

**Chauã Coluene Queirolo**

**Reconhecimento Facial 3D usando Simulated Annealing  
e a Medida de Interpenetração de Superfícies**

**CURITIBA**

**2009**

**Chauã Coluene Queirolo**

**Reconhecimento Facial 3D usando Simulated Annealing  
e a Medida de Interpenetração de Superfícies**

Dissertação apresentada como requisito parcial para  
obtenção do grau de Mestre. Programa de Pós-  
Graduação em Informática, Setor de Ciências Exa-  
tas, Universidade Federal do Paraná.

Orientadores: Olga Regina Pereira Bellon  
Luciano Silva

**CURITIBA**

**2009**

# SUMÁRIO

<b>Lista de Figuras .....</b>	<b>iii</b>
<b>Lista de Tabelas .....</b>	<b>v</b>
<b>Lista de Abreviaturas .....</b>	<b>vii</b>
<b>Abstract .....</b>	<b>viii</b>
<b>Resumo .....</b>	<b>ix</b>
<b>1 Introdução .....</b>	<b>1</b>
<b>2 Revisão Literária .....</b>	<b>5</b>
<b>3 Base de Dados .....</b>	<b>10</b>
<b>4 Pré-processamento .....</b>	<b>14</b>
<b>5 Alinhamento de Faces 3D .....</b>	<b>16</b>
5.1 <i>Surface Interpenetration Measure (SIM)</i> .....	17
5.2 <i>Simulated Annealing (SA)</i> .....	19
5.3 <i>Simulated Annealing</i> Modificado .....	22
5.4 Lista de parâmetros .....	23
<b>6 Autenticação de Imagens Faciais 3D .....</b>	<b>25</b>
6.1 Modelo de Avaliação Hierárquico .....	27
6.2 Modelo de Avaliação Hierárquico Estendido .....	28

<b>7</b>	<b>Resultados Experimentais .....</b>	<b>30</b>
7.1	Desempenho do SA para o Alinhamento de Faces 3D .....	31
7.2	Experimento I: Verificação .....	32
7.2.1	Análise dos Buracos na Região dos Olhos .....	33
7.2.2	Análise das Expressões Faciais .....	33
7.2.3	Análise do Modelo Hierárquico .....	35
7.3	Experimento II: Identificação .....	36
7.4	Experimento III: <i>Face Recognition Grand Challenge</i> (FRGC) ROC III .....	38
7.5	Experimento IV: Modelo de Avaliação Hierárquico .....	38
7.6	Comparação com Outros Métodos.....	40
7.7	Experimento V: Melhoria na Taxa de Reconhecimento .....	42
<b>8</b>	<b>Considerações Finais .....</b>	<b>44</b>
	<b>Referências Bibliográficas .....</b>	<b>46</b>

## LISTA DE FIGURAS

Figura 1.1	Etapas principais para a autenticação de faces 3D. ....	3
Figura 3.1	Exemplos de imagens da base de dados FRGC v2. ....	10
Figura 3.2	Exemplos de imagens com artefatos: (a) imagens esticadas/deformadas, (b) ausência do nariz, (c) buracos ao redor do nariz, e (d) “ondas” ao redor da boca. ....	11
Figura 3.3	Exemplo de imagens classificadas como neutras na base FRGC v2. ....	12
Figura 3.4	Indivíduo com a classificação incorreta: (a)-(b) imagens 2D e 3D do “impostor”, respectivamente, (c)-(d) imagens 2D e 3D imagens do outro indivíduo, respectivamente. ....	13
Figura 4.1	Deteção dos pontos característicos: (a) localização dos seis pontos, (b) deteção incorreta dos pontos do nariz, e (c) deteção incorreta dos cantos dos olhos. ....	15
Figura 4.2	Regiões segmentadas da mesma face: (1) região circular e (2) elíptica ao redor do nariz, (3) parte superior da cabeça, e (4) face inteira com os pontos característicos detectados. ....	15
Figura 5.1	Ponto interpenetrado $p$ de $A$ em relação a $B$ [51]. ....	17
Figura 5.2	Etapas principais do SA: (1) solução inicial, (2) alinhamento grosseiro usando MSAC, e (3) alinhamento fino usando a SIM. ....	21

Figura 5.3	Setores utilizados pelo SA modificado: (a) imagem original, (b) nove setores utilizados, e (c) regiões consideradas invariantes em destaque. ....	23
Figura 6.1	Diagrama do Modelo de Avaliação Hierárquico [16]. ....	28
Figura 7.1	Desempenho para o experimento “All vs. All”. ....	33
Figura 7.2	Resultados para base FRGC usando faces com expressões faciais. ....	35
Figura 7.3	Conjunto de curvas CMC: Level 0 vs. Others, Level 1 vs. Others, First vs. Others, e All vs. All. ....	37
Figura 7.4	Desempenho no experimento FRGC ROC III: (a) curva <i>False Acceptance Rate</i> (FAR), e (b) curva <i>Cumulative Match Characteristic</i> (CMC). ....	39

## LISTA DE TABELAS

Tabela 2.1	Sumário dos principais trabalhos relacionados. ....	9
Tabela 3.1	Classificação das imagens da base FRGC v2. ....	12
Tabela 3.2	Classificação da base FRGC v2 por expressões faciais. ....	13
Tabela 5.1	Relação entre o processo <i>Annealing</i> real e o <i>Simulated Annealing</i> [15]. ..	19
Tabela 5.2	Lista dos parâmetros utilizados. ....	24
Tabela 6.1	Taxa de verificação para cada região da face, com FAR de 0%. ....	25
Tabela 6.2	Taxa de verificação usando a regra da soma, com FAR de 0%. ....	26
Tabela 6.3	Medidas calculadas em cada nível da hierarquia. ....	29
Tabela 7.1	Classificação dos conjuntos de imagens da base FRGC v2. [47] ....	31
Tabela 7.2	Tempo médio de execução do SA para alinhar duas faces 3D. ....	31
Tabela 7.3	Experimento I: Taxa de verificação usando FAR de 0% e FAR de 0,1%. ..	32
Tabela 7.4	Resultados para base FRGC usando faces com expressões faciais. ....	34
Tabela 7.5	Taxa de verificação em cada nível no Modelo de Avaliação Hierárquico. .	35

Tabela 7.6	Taxa de verificação em cada nível no Modelo de Hierárquico Estendido. .	36
Tabela 7.7	Experimento II: Taxa de reconhecimento <i>rank-one</i> . . . . .	37
Tabela 7.8	Experimento III: Resultados para FRGC ROC III. . . . .	38
Tabela 7.9	Experimento IV: Taxa de verificação usando FAR 0%. . . . .	40
Tabela 7.10	Resultados de verificação usando a base All vs. All, com 0,1% de FAR. . .	40
Tabela 7.11	Resultados de identificação usando conjunto First vs. Others. . . . .	41
Tabela 7.12	Resultados de verificação para o experimento FRGC ROC III, com 0,1% de FAR. . . . .	41
Tabela 7.13	Resultados do experimento ROC III em conjuntos diferentes, com 0,1% de FAR. . . . .	41
Tabela 7.14	Experimento V: Melhoria na taxa de reconhecimento. . . . .	43



## LISTA DE ABREVIATURAS

<b>AFM</b>	<i>Annotated Face Model</i>
<b>AG</b>	<i>Algoritmos Genéticos</i>
<b>ARMS</b>	<i>Adaptive Rigid Multi-Region Selection</i>
<b>CMC</b>	<i>Cumulative Match Characteristic</i>
<b>FAR</b>	<i>False Acceptance Rate</i>
<b>FRGC</b>	<i>Face Recognition Grand Challenge</i>
<b>HGM</b>	<i>Hierarchical Graph Matching</i>
<b>ICP</b>	<i>Iterative Closest Point</i>
<b>LDA</b>	<i>Linear Discriminant Analysis</i>
<b>MSAC</b>	<i>M-estimator Sample Consensus</i>
<b>MSE</b>	<i>Mean Squared Error</i>
<b>PCA</b>	<i>Principal Component Analysis</i>
<b>SA</b>	<i>Simulated Annealing</i>
<b>SIFT</b>	<i>Scale-Invariant Feature Transform</i>
<b>SIM</b>	<i>Surface Interpenetration Measure</i>

## ABSTRACT

This work presents a novel automatic framework to perform 3D face recognition. The proposed method uses a Simulated Annealing-based approach (SA) for range image registration with the *Surface Interpenetration Measure* (SIM) as similarity measure in order to match two face images. The authentication score is obtained by combining the SIM values corresponding to the matching of four different face regions: (1) circular and (2) elliptical areas around the nose, (3) forehead and (4) the entire face region. Then, a modified SA approach is proposed taking advantage of invariant face regions to better handle facial expressions. Comprehensive experiments were performed on the FRGC v2 database, which is the largest available database of 3D face images composed of 4,007 images with different facial expressions. The experiments simulated both verification and identification systems and the results were compared to those reported by state-of-the-art works. By using all the images in the database, a verification rate of 96.5% was achieved at a *False Acceptance Rate* (FAR) of 0.1%. In the identification scenario, a rank-one accuracy of 98.4% was achieved. To our best knowledge, this is the highest rank-one score ever achieved for the FRGC v2 database when compared to results published in the literature.

**Keywords:** 3D Face Recognition, FRGC Database, *Surface Interpenetration Measure* (SIM), Simulated Annealing, Range Image Registration.

## RESUMO

Este trabalho apresenta um novo sistema automático para reconhecimento facial usando imagens 3D. O método proposto é baseado no algoritmo *Simulated Annealing* (SA) para o alinhamento de imagens de profundidade, sendo usada a *Surface Interpenetration Measure* (SIM) como medida de similaridade entre duas imagens. A medida de autenticação é obtida combinando a SIM calculada para quatro regiões da face: áreas circular e elíptica ao redor do nariz, parte superior da face e a face inteira. Além disso, uma abordagem modificada do SA é usada para minimizar os efeitos decorrentes das expressões faciais durante o alinhamento. Uma série de experimentos foram realizados na base de dados *Face Recognition Grand Challenge* (FRGC) v2, que é a maior base de faces 3D disponível atualmente, composta por 4.007 imagens com diferentes expressões faciais. Os experimentos realizados simularam sistemas de verificação e identificação, e os resultados obtidos foram comparados com os trabalhos estado-da-arte presentes na literatura. Usando todas as imagens da base FRGC v2 foi obtida uma taxa de verificação de 96,5%, com uma Taxa de Falsa Aceitação (*False Acceptance Rate* - FAR) de 0,1%. No cenário de identificação foi obtido um *rank-one* de 98,4%. Pelo nosso conhecimento, estes são os melhores resultados de identificação já apresentados usando a base FRGC v2, quando comparados com outros resultados apresentados na literatura.

**Palavras-chave:** Reconhecimento facial 3D, FRGC Database, *Surface Interpenetration Measure* (SIM), *Simulated Annealing*, Registro de imagens de profundidade.

# 1 INTRODUÇÃO

Sistemas de reconhecimento facial possuem aplicações em diversas áreas, tais como segurança, entretenimento, vigilância, controle de acesso [29]. Nas últimas décadas, os estudos nessa área tiveram um grande avanço com a utilização de imagens 2D. No entanto, a utilização de imagens faciais 2D possui limitações decorrentes da mudança de pose, iluminação e expressões faciais [54]. Uma maneira de lidar com essas limitações é a utilização da informação de profundidade da face. Na última década o reconhecimento facial 3D ganhou destaque devido aos avanços dos sensores para aquisição de imagens 3D. No entanto, a utilização de imagens 3D também possui algumas limitações, como a presença de ruídos e a dificuldade na aquisição [8].

Existem duas abordagens principais para trabalhar com a informação facial 3D, dependendo do formato dos dados [24]: (1) imagem de profundidade (*range image*) e (2) nuvem de pontos. Para imagens de profundidade, uma técnica popular é a utilização de métodos estatísticos, como a Análise das Componentes Principais (*Principal Component Analysis (PCA)*) [38]. Para aplicar o PCA, primeiramente as faces 3D são normalizadas e em seguida projetadas como imagens 2D. Uma das vantagens desta abordagem é que os métodos podem ser facilmente combinados com as imagens de textura para aumentar a robustez do sistema [8,25,43].

Para nuvens de pontos, as técnicas de registro de imagens são geralmente usadas para alinhar os dois conjuntos de pontos. O método *Iterative Closest Point* (ICP) [6, 11], ou alguma de suas variantes [23,49], é normalmente utilizado para realizar essa tarefa. O Erro Quadrático Médio (*Mean Squared Error (MSE)*), minimizado durante o ICP, é então usado como base para definir a similaridade entre duas imagens faciais [10, 12, 21, 35]. Essa abordagem também pode ser combinada com técnicas de deformação para modelar as expressões faciais, e com isso minimizar o seu efeito no reconhecimento [26, 30, 34].

O desempenho dos sistemas de reconhecimento facial 2D e 3D pode ser avaliado no *Face Recognition Grand Challenge* (FRGC) [44]. O FRGC é uma plataforma internacional composta por seis conjuntos de problemas e suas taxas de desempenho esperadas. O terceiro experimento, *i.e.*, ROC III, se refere ao problema de reconhecimento facial 3D que utiliza uma

base composta por 4.007 imagens de 466 pessoas com diferentes expressões faciais. Neste experimento o objetivo é atingir uma taxa de reconhecimento de 98,0% com uma Taxa de Falsa Aceitação (*False Acceptance Rate* (FAR)) de 0,1% [44].

A motivação deste trabalho é explorar a utilização da informação de profundidade da face para o reconhecimento. Dadas duas imagens faciais de entrada, o objetivo do sistema é diferenciar se ambas as faces pertencem ou não a um mesmo indivíduo. Para validar a precisão de tal sistema duas medidas são utilizadas: (1) FAR, e (2) *rank-one*. O FAR é empregado para avaliar sistemas biométricos no modo de verificação. Neste modo, duas imagens faciais são comparadas e o sistema responde se pertencem ou não à mesma pessoa. O *rank one* é utilizado para avaliar sistemas de identificação, o qual retorna a identidade da pessoa, *i.e.*, retorna a face com maior similaridade com a face de entrada [7].

Neste trabalho apresentamos um sistema para reconhecimento facial que tem como entrada apenas a informação de profundidade das faces. Esta abordagem pode ser usada para verificação e identificação, e possui quatro estágios principais (ver Fig. 1.1): (1) aquisição das imagens, (2) pré-processamento, (3) alinhamento, e (4) autenticação.

Inicialmente, a aquisição das imagens faciais 3D pode ser realizada usando diferentes técnicas, como por exemplo, sensores de luz estruturada ou sensores *laser* [22]. As imagens utilizadas nesse trabalhos foram adquiridas por um sensor *laser*, que produz uma imagem de profundidade do objeto capturado. Na etapa de pré-processamento, a imagem de entrada é suavizada e a região da face é detectada e segmentada automaticamente. Então, alguns pontos característicos, como o canto dos olhos e a ponta do nariz, são detectados para serem utilizados durante o processo de alinhamento [41, 42].

Com isso, cada uma das regiões segmentadas é alinhada com a sua região correspondente de uma face que foi previamente cadastrada em um banco de dados. O alinhamento é realizado usando uma abordagem baseada no algoritmo *Simulated Annealing* (SA) que utiliza duas medidas robustas de avaliação: (1) *M-estimator Sample Consensus* (MSAC) [53], e a (2) *Surface Interpenetration Measure* (SIM) [51]. Além disso, também é usada uma abordagem modificada do SA para lidar com as expressões faciais [45].

Finalmente, a medida de autenticação é calculada com base no valor da SIM obtida para o registro de cada região. A medida resultante da soma pode ser usada para: (1) verificar se duas faces pertencem ou não à mesma pessoa, e (2) recuperar a identidade de uma pessoa. Em trabalhos anteriores a SIM mostrou ser uma medida discriminatória quando aplicada no reconhecimento facial 3D [4, 5, 45, 46, 47].

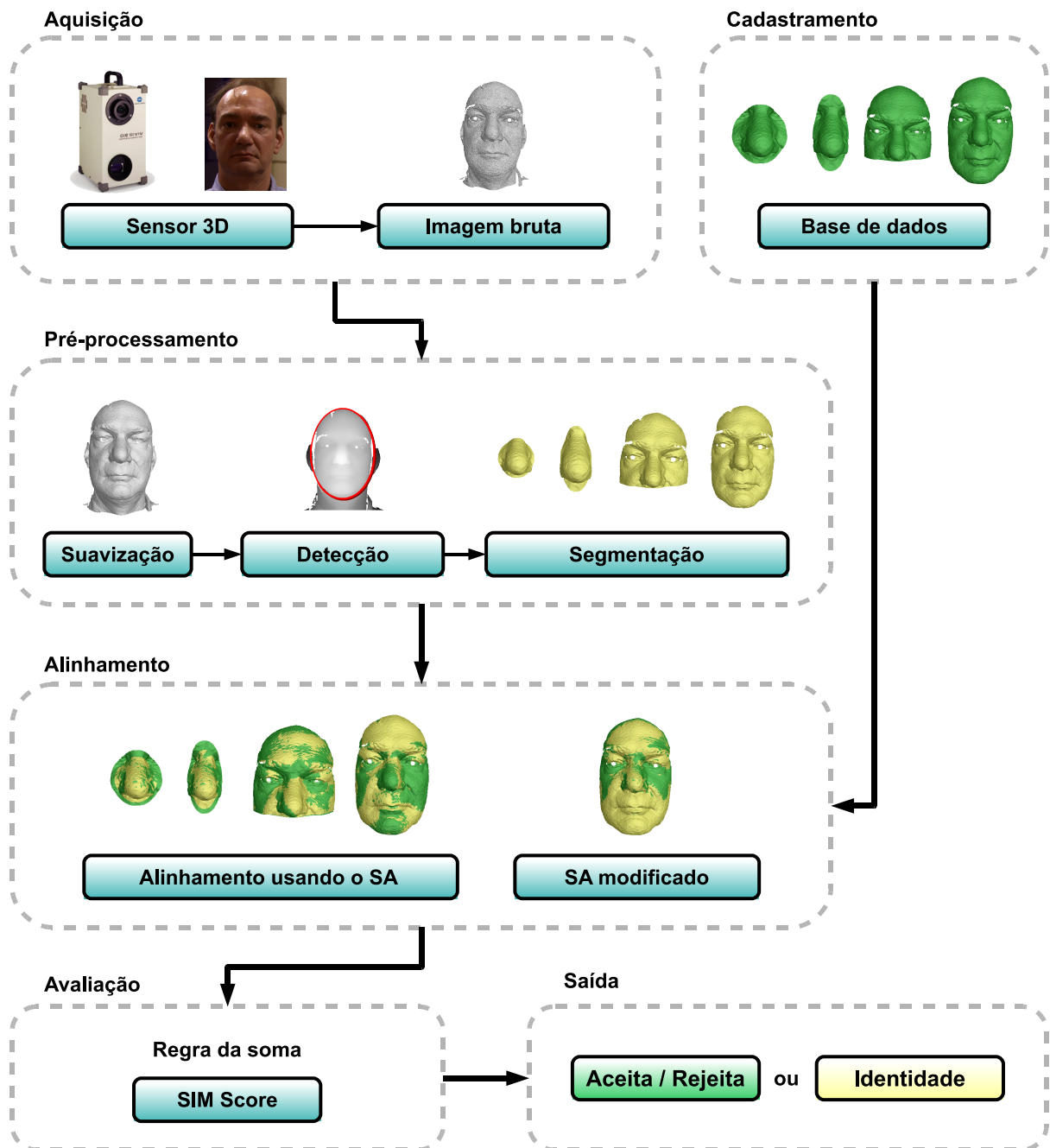


Figura 1.1: Etapas principais para a autenticação de faces 3D.

Como pode ser observado em diversos trabalhos na literatura [7, 29], não existe um sistema biométrico que seja 100% ideal. Além disso, os estudos com a utilização da informação 3D em problemas de visão computacional têm sido bastante explorados. A idéia deste trabalho é propor uma nova medida para realizar a comparação entre duas faces focando em um sistema biométrico robusto. Tal sistema visa a obtenção de 0% de falsa rejeição, e ao mesmo tempo 100% de *rank-one*.

Uma das contribuições principais deste trabalho é propor a SIM como medida de similaridade entre duas imagens faciais 3D [47]. Ao contrário dos trabalhos presentes na literatura que focam no alinhamento usando o ICP (guiado pelo MSE), este trabalho se baseia no fato de que Silva *et al.* [51] provaram que o MSE permite a convergência para mínimos locais, até mesmo quando são utilizadas abordagens aprimoradas do ICP [23, 49]. Os autores também sugerem que o MSE é uma boa medida para iniciar o processo do alinhamento, sendo a SIM mais apropriada nas etapas finais para garantir a qualidade do alinhamento. Outras contribuições deste trabalho que podem ser destacadas é a classificação da base FRGC v2 em conjuntos controlados, sendo a base de segmentada de acordo com o tipo de expressão facial e a sua intensidade, realização de diferentes experimentos na base FRGC v2 em diferentes contextos, análise de diferentes abordagens para autenticação de faces, extensão do modelo hierárquico, e comparação com outros trabalhos presentes na literatura [45, 46].

Os experimentos foram realizados usando sub-bases controladas, da base FRGC v2, para avaliação de desempenho do sistema proposto. Além disso, apresentamos uma análise comparativa de resultados com os trabalhos relacionados mais relevantes nessa área [2, 13, 21, 30, 33, 39, 40]. Esta análise é uma das mais completas já descritas na literatura sobre reconhecimento facial 3D, com base em nossos conhecimentos. Faltemier *et al.* [21] também apresenta um conjunto completo de resultados experimentais.

Este trabalho está dividido como segue. A Seção 2 apresenta uma discussão sobre os principais trabalhos relacionados. A Seção 3 apresenta detalhes sobre a base de dados utilizada e a Seção 4 descreve a etapa de pré-processamento. A Seção 5 apresenta o método para alinhamento de faces 3D e é seguida do processo de autenticação na Seção 6. A Seção 7 apresenta os resultados experimentais e a Seção 8 conclui com as considerações finais.

## 2 REVISÃO LITERÁRIA

Esta seção apresenta uma revisão dos trabalhos mais relevantes que abordam o reconhecimento facial 3D. Esses trabalhos, em sua maioria, calculam a medida de similaridade entre duas imagens faciais com base no MSE minimizado durante a etapa de alinhamento, algumas vezes combinado com outras medidas.

Chang *et al.* [10] descrevem um método chamado *Adaptive Rigid Multi-Region Selection* (ARMS). O ARMS usa três regiões de sobreposição ao redor do nariz: (1) círculo, (2) elipse, e (3) o nariz propriamente dito. A região do nariz foi escolhida por ser invariante à presença de expressões faciais. Então, cada região é alinhada com sua correspondente usando o algoritmo ICP [6]. O ICP é um algoritmo que dada duas superfícies, calcula a melhor transformação que minimize a distância entre os pontos correspondentes. O MSE resultante de cada alinhamento é combinado usando a regra do produto, a qual é usada como medida de similaridade entre duas faces 3D. Os experimentos foram realizados usando um super conjunto de imagens da base FRGC, composto por 4.485 imagens faciais 3D. Quando utilizadas apenas as faces com expressão neutra, o método obtém uma taxa de *rank-one* de 97,1%. A taxa de *rank-one* obtida utilizando faces com expressão foi de 87,1%.

Lu *et al.* [35] apresentam uma abordagem multimodal para o reconhecimento facial composta por dois módulos principais, um para imagens 3D e outro para imagens 2D. O método *Linear Discriminant Analysis* (LDA) [3] é usado para verificar a similaridade das imagens faciais 2D. O LDA é um método estatístico, da mesma família que o PCA, que separa um grupo de objetos em classes mutualmente exclusivas usando como base um conjunto de características. Para as imagens 3D é empregado o ICP, sendo o MSE usado como medida de similaridade. Os experimentos foram realizados em uma base de dados privada com 200 modelos faciais 3D usados como *gallery*, e 598 imagens usadas como *probe*. A medida de similaridade é calculada usando a regra da soma ponderada entre as medidas obtidas pelo LDA e pelo ICP. A abordagem proposta obteve *rank-one* de 90% e os autores verificaram que a presença de expressões faciais impediu a identificação correta de todas as imagens.



Maurer *et al.* [39] propõem um método que combina a utilização de imagens faciais 2D e 3D. Para o reconhecimento facial 2D é usado um software comercial, porém, os autores sugerem que essa etapa pode ser realizada por qualquer outra abordagem. Para o reconhecimento facial 3D é utilizado o algoritmo ICP, e o mapa de distâncias é usado como medida de similaridade. As medidas computadas para as imagens 2D e 3D são combinadas usando uma soma ponderada. Na FRGC v2 foi obtida uma taxa de verificação de 87% com uma FAR de 0,1%, considerando apenas a informação de profundidade. Os autores classificaram as principais fontes causadoras de erros em três categorias: (1) imagens distorcidas, (2) imagens 2D e 3D não sincronizadas, e (3) outros, *e.g.*, cabelo cobrindo a face.

Husken *et al.* [27] apresentam a fusão dos métodos *Hierarchical Graph Matching* (HGM) 2D e HGM 3D para realizar o reconhecimento facial. Essa abordagem usa um grafo que armazena informações sobre a textura da imagem e as posições dos pontos característicos, *e.g.*, nariz, olhos e boca. Então, o grafo é adaptado automaticamente à face usando esses pontos característicos, e a medida de similaridade é calculada a partir da distância entre dois grafos. Nos experimentos usando apenas as imagens 3D da FRGC v2, foi obtida uma taxa de verificação de 86,9%, com uma FAR de 0,1%.

Lin *et al.* [33] descrevem um método que combina a soma de características invariantes da face para sua classificação. Dez áreas retangulares são selecionadas da imagem de entrada para que seja extraído um conjunto de características. As métricas do alinhamento de cada região são combinadas usando a regra da soma ou a regra ponderada da soma, otimizada pelo LDA. Usando o LDA, para definir esses pesos, é possível alcançar maiores taxas de reconhecimento. Os experimentos foram realizados na base FRGC v2, e no experimento ROC III foi obtida uma taxa de verificação de 90%, com uma FAR de 0,1%.

Bronstein *et al.* [9] apresentam um método multimodal usando uma representação da face invariante às expressões faciais. Eles assumem que as expressões podem ser modeladas como isomerias da superfície tridimensional para criar um modelo facial canônico. Uma vez criados os modelos, o alinhamento das superfícies é realizado usando momentos de alta-ordem, que é um procedimento simples e eficiente para o registro de imagens [19]. Embora essa abordagem não sofra efeitos das expressões faciais, a representação canônica das faces pode causar a perda de algumas características importantes, como o formato do nariz e os cantos dos olhos. Os experimentos foram realizados em uma base de imagens privada composta por 220 imagens de 30 pessoas, com diferentes expressões faciais. Os resultados mostraram que o registro de superfícies canônicas é mais discriminante do que o registro de superfícies rígidas.

Mian *et al.* [40] descrevem um sistema de reconhecimento facial multimodal 2D+3D. Inicialmente, a pose da face 3D é corrigida e normalizada automaticamente com a sua imagem 2D correspondente. O algoritmo *Scale-Invariant Feature Transform* (SIFT) é usado nas imagens faciais 2D para detectar pontos característicos locais. A SIFT mostrou ser mais robusta com relação à iluminação e à presença de expressões faciais, quando comparada com o PCA. A similaridade entre as imagens 2D é determinada pela distância Euclideana dos descritores das características faciais. As imagens faciais 3D são segmentadas em duas regiões: (1) olhos e testa, e (2) região do nariz. Cada região é alinhada com as suas correspondentes usando o algoritmo ICP, e a medida de similaridade é calculada combinando o resultado de cada alinhamento. Na base FRGC v2 foi obtida uma taxa de verificação de 86,6% com uma FAR de 0,1%, quando utilizadas apenas as imagens 3D. Comparando as faces neutras com todas as outras imagens restantes da base, a taxa de verificação aumenta para 98,5%, com uma FAR de 0,1%.

Cook *et al.* [13] aplicam filtros *Log-Gabor* para o reconhecimento facial 3D. Após a aplicação do filtro, a imagem facial é segmentada em 49 regiões quadradas e decomposta por três filtros de escala, totalizando 147 subregiões. O PCA é aplicado em cada subregião para representar a face com 147 características. A distância *Mahalanobis-Cossine* é empregada para calcular a distância entre dois vetores de características. Nos resultados experimentais, o melhor desempenho foi obtido usando apenas as 75 subregiões centrais da face. Na base FRGC v2 foi obtida uma taxa de verificação de 92,3% com 0,1% de FAR e um *rank-one* de 96,2%.

Kakadiaris *et al.* [30] apresentam um processo para autenticação de faces 3D completamente automático usando *Annotated Face Model* (AFM) para lidar com as expressões faciais. Inicialmente um AFM é criado para ser deformado na etapa de alinhamento permitindo assim a criação de um meta-dado para a autenticação. Nas etapas iniciais do processo, a imagem de entrada é processada para a remoção de artefatos e o preenchimento de buracos. Uma amostra dos dados é criada para melhorar o desempenho do método. Os dados de entrada são alinhados ao modelo AFM usando um método de alinhamento composto das seguintes etapas: (1) *Spin Images*, para definir a correspondência inicial entre o modelo AFM e os dados de entrada; (2) ICP, usado para realizar a transformação rígida; e o (3) SA, usado para assegurar um alinhamento mais preciso. Depois dessa etapa, é obtido um modelo deformável, o qual é convertido para um modelo geométrico e para um mapa normal. Cada uma dessas imagens é então tratada separadamente usando duas transformações *wavelets*, *Haar* e *Pyramid*. A autenticação é calculada no domínio *wavelet* sendo definidas diferentes métricas de distância para cada *wavelet*. No experimento ROC III, da base FRGC v2, os autores reportaram uma taxa de verificação de 97,0%, com uma FAR de 0,1%. Além disso, eles reportaram um *rank-one* de 97,0%.

Al-Osaimi *et al.* [2] apresentam uma abordagem para reconhecimento facial 3D usando modelos deformáveis. Para isso, um subespaço PCA é criado com base nas deformações das expressões faciais. Inicialmente, a etapa de treino requer imagens neutras e não-neutras de cada indivíduo. As imagens são segmentadas e alinhadas com o ICP para verificar as superfícies residuais. Então, um subespaço PCA é criado com base nesses resíduos, que conseqüentemente consegue modelar uma deformação de expressão facial genérica. A similaridade é calculada comparando as faces 3D deformadas. Os experimentos foram conduzidos na base FRGC v2 em diferentes subconjuntos de imagens. Foi obtida uma taxa de verificação de 98,35% e 97,7%, com uma FAR de 0,1%, nos conjuntos de faces neutras e não-neutras, respectivamente. No experimento ROC III, foi obtida uma taxa de verificação de 94,1%, com uma FAR de 0,1%.

Faltemier *et al.* [21] descrevem um sistema que consiste na fusão de 28 regiões esféricas extraídas da face. Essas regiões são então combinadas usando o algoritmo ICP. O MSE resultante do alinhamento de cada uma das regiões foi combinado usando diferentes abordagens, e os melhores resultados foram obtidos usando os métodos *Consensus Voting* e *Borda Counting*. Para tornar o sistema robusto em relação às variações faciais, que podem ocorrer em pequenas regiões da face, são usadas diferentes regiões para o alinhamento. Na base FRGC v2 a taxa de verificação obtida foi de 93,2%, com uma FAR de 0,1%, e o *rank-one* de 97,2%.

Um *survey* de trabalhos relacionados com reconhecimento facial 3D e multimodal 2D+3D é apresentado em [1, 8]. A Tabela 2.1 apresenta um sumário dos trabalhos apresentados nessa seção. Como pode ser observado, a maioria dos trabalhos utiliza com sucesso o algoritmo ICP para o reconhecimento facial 3D [10, 21, 39, 40], sendo o MSE usado como medida de similaridade, algumas vezes combinado com outras medidas. Neste trabalho, nós apresentamos uma nova abordagem usando a SIM para obter resultados mais precisos.

Tabela 2.1: Sumário dos principais trabalhos relacionados.

	<b>Modalidade</b>	<b>Bases</b>	<b>Método</b>
Chang <i>et al.</i> [10]	3D	FRGC v1+v2	ICP
Lu <i>et al.</i> [35]	2D+3D	Própria	LDA (imagens 2D) + ICP (imagens 3D).
Maurer <i>et al.</i> [39]	2D+3D	FRGC v2	ICP
Husken <i>et al.</i> [27]	2D+3D	FRGC v2	HGM 2D + HGM 3D.
Lin <i>et al.</i> [33]	3D	FRGC v2	Soma de características invariantes da face
Bronstein <i>et al.</i> [9]	2D+3D	Própria	Momentos de alta-ordem
Mian <i>et al.</i> [40]	2D+3D	FRGC v2	SIFT (imagens 2D) + ICP (imagens 3D)
Cook <i>et al.</i> [13]	3D	FRGC v2	<i>Log-Gabor</i>
Kakadiaris <i>et al.</i> [30]	3D	FRGC v2	Modelos deformáveis + <i>Wavelets</i>
Al-Osaimi <i>et al.</i> [2]	3D	FRGC v2	ICP
Faltemier <i>et al.</i> [21]	3D	FRGC v2	ICP
Queirolo <i>et al.</i> [47]	3D	FRGC v2	SA + SIM

### 3 BASE DE DADOS

Nesta seção será discutida a primeira etapa do sistema de reconhecimento facial proposto, como apresentado na Fig. 1.1. Neste trabalho, as imagens utilizadas tanto na aquisição quanto no cadastramento pertencem à base FRGC v2. Esta seção apresenta uma discussão sobre as características das imagens presentes nessa base.

Nos experimentos foi utilizada a base de dados FRGC v2, que é a maior base de imagens faciais 3D disponível atualmente. Essa base é usada como referência para avaliação de métodos de reconhecimento facial [44]. A base FRGC v2 possui 4.007 imagens de 466 pessoas diferentes. Cada pessoa possui no máximo 22 imagens, sendo que 56 indivíduos possuem apenas uma imagem. Todas as imagens foram adquiridas com um sensor *laser Minolta Vivid 910* e têm resolução 640x480. A Fig. 3.1 apresenta algumas das imagens presentes na base FRGC v2. As legendas das figuras representam o nome do arquivo de cada imagem, composto por duas partes com base no delimitador “d”: (1) identificação do indivíduo, e (2) número da imagem.

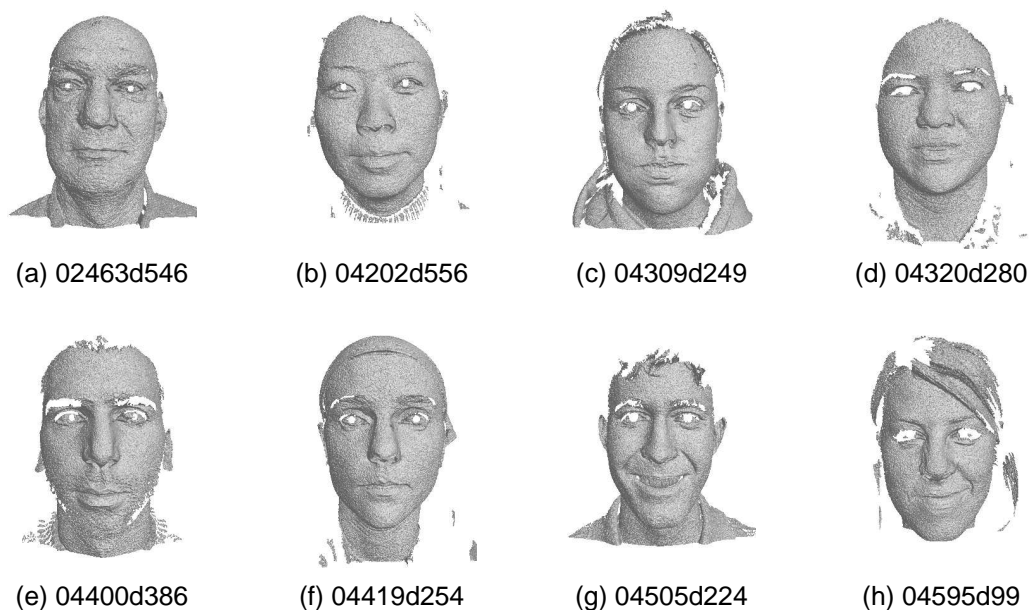


Figura 3.1: Exemplos de imagens da base de dados FRGC v2.

As faces possuem pose frontal e expressões faciais variadas. Os tipos de expressões mais comuns na base são: neutro, alegre, triste, nojo e surpresa. Algumas imagens também apresentam artefatos que podem ser classificados como: (1) imagens distorcidas, (2) ausência do nariz, (3) buracos ao redor do nariz, ou (4) “ondas” ao redor da boca. Exemplos desses tipos de imagens são apresentados na Fig. 3.2, e as legendas indicam a sua identificação na base.

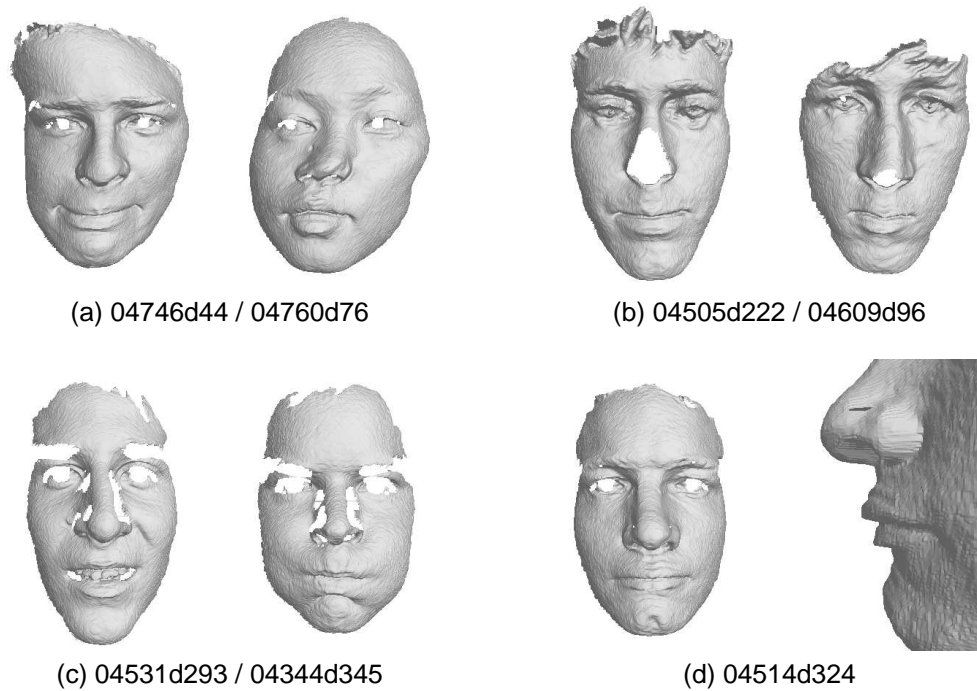


Figura 3.2: Exemplos de imagens com artefatos: (a) imagens esticadas/deformadas, (b) ausência do nariz, (c) buracos ao redor do nariz, e (d) “ondas” ao redor da boca.

O primeiro tipo de artefato é ocasionado pela movimentação da pessoa durante o processo de aquisição das imagens. O segundo está relacionado com a distância entre a pessoa e o aparelho de aquisição; se a face estiver muito próxima, as regiões mais próximas ao sensor podem não ser capturadas devido à limitação da distância focal. Outro problema ocorre porque os sensores *lasers* não conseguem mapear as regiões transparentes e com alta especularidade, resultando em buracos nas imagens capturadas, *e.g.*, região dos olhos e ao redor do nariz. Bowyer *et al.* [8] apresentam um estudo onde foram identificadas diversas situações que podem afetar a qualidade da imagem de profundidade, como por exemplo, a iluminação do ambiente. Algumas dessas particularidades da base FRGC v2 também são descritas em [39].

Como o objetivo é realizar uma avaliação detalhada da nossa abordagem para o reconhecimento facial 3D, a base FRGC v2 foi classificada de acordo com as expressões faciais e o nível de ruído (relacionado com a presença de artefatos). Realizamos uma nova classificação das expressões faciais porque a classificação disponível pela FRGC contém algumas incon-

sistências. A Fig. 3.3 apresenta alguns exemplos de faces que foram classificadas como neutras, mas que de fato possuem algum tipo de expressão. Uma razão provável é que a classificação foi realizada de acordo com a expressão solicitada, e não com a expressão capturada de fato. Neste trabalho a FRGC v2 foi segmentada em diferentes conjuntos, como apresentado na Tabela 3.1.

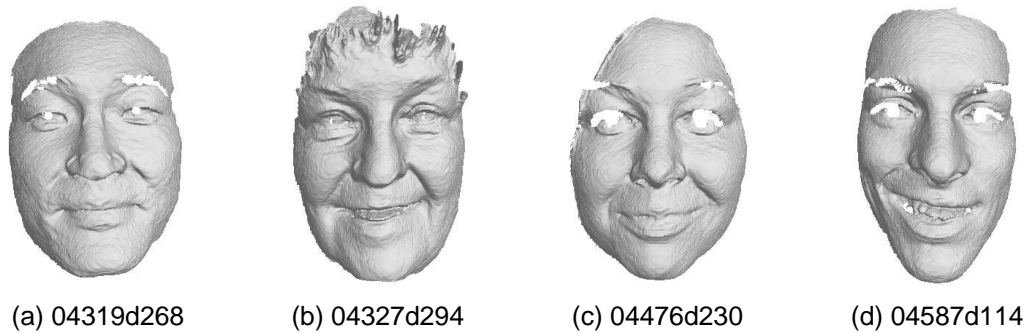


Figura 3.3: Exemplo de imagens classificadas como neutras na base FRGC v2.

Tabela 3.1: Classificação das imagens da base FRGC v2.

Conjuntos de imagens	Número de imagens
Pouca expressão com artefatos	247
Pouca expressão sem artefatos	455
Neutro sem artefatos	933
Neutro com/sem artefatos	2.182
Todas	4.007

Para verificar a influência dos buracos na região dos olhos no reconhecimento, a base FRGC v2 foi segmentada em três conjuntos: (1) sem buracos, com 51 imagens, (2) buracos pequenos, com 1.300 imagens, e (3) buracos grandes, com 2.656 imagens. Esse efeito é decorrente do fato do olho ser uma região translúcida, o que faz com que o feixe seja refratado ao invés de ser refletido. Por essa razão, o sensor *laser* não consegue capturar adequadamente a reflexão do feixe, impedindo a geração da informação 3D dessa região. Nós observamos que quase todas as imagens da base apresentam esse efeito e por isso nós não consideramos os buracos nos olhos como sendo um artefato na classificação da Tabela 3.1. Experimentos mais detalhados para verificar a sua influência no reconhecimento são apresentados na Seção 7.2.1.

Além disso, as imagens da base FRGC v2 também foram classificadas de acordo com o tipo de expressão facial, usando como base as seis emoções básicas propostas por [18]. As seguintes classes foram definidas: neutra, neutra com boca aberta (B.A.), alegre, surpresa, triste,

franzir, nojo e “bochechudo” (*puffy cheek*). Com exceção das expressões neutra e “bochechudo”, as classes restantes também foram organizadas de acordo com sua intensidade. A Tabela 3.2 apresenta o número de imagens de cada conjunto.

Tabela 3.2: Classificação da base FRGC v2 por expressões faciais.

Conjuntos de imagens	Número de imagens
Pouco alegre	208
Muito alegre	249
Pouco surpresa	56
Muito surpresa	172
Pouco triste	82
Muito triste	57
Pouco boca aberta	64
Muito boca aberta	212
Pouco franzir	32
Muito franzir	52
Pouco nojo	13
Muito nojo	59
“Bochechudo”	150
Outros	66

Na base FRGC v2 foi observado que dois indivíduos, embora possuam identificações diferentes, são na verdade a mesma pessoa (ver Fig. 3.4). Essa informação foi confirmada pela organização do FRGC. O número de identificação do impostor na base é 04643, e o número de identificação do outro indivíduo é 04783. Esse rótulo foi corrigido para os nossos experimentos.

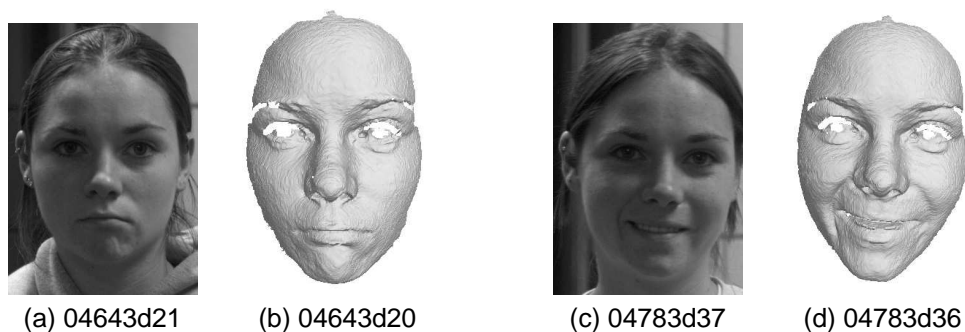


Figura 3.4: Indivíduo com a classificação incorreta: (a)-(b) imagens 2D e 3D do “impostor”, respectivamente, (c)-(d) imagens 2D e 3D imagens do outro indivíduo, respectivamente.



## 4 PRÉ-PROCESSAMENTO

Esta seção apresenta uma discussão sobre como é realizada a segmentação da imagem de entrada e da extração da face. Uma vez adquiridas as imagens faciais 3D, é necessário extrair a região de interesse (*i.e.* região da face) da imagem de entrada, como apresentado na Fig. 1.1. As imagens de entrada possuem algumas áreas que podem interferir na etapa de alinhamento das faces 3D, tais como cabelo, pescoço e orelhas (ver Fig.3.1). Inicialmente a face é suavizada com o filtro da mediana, seguido pelo processo de segmentação que utiliza a nossa própria abordagem baseada em mapas de profundidade. Essa abordagem possui duas etapas principais: (1) localização das regiões homogêneas da imagem, combinando informações de borda e agrupamento de regiões, e (2) identificação das regiões candidatas à região da face, usando um método de detecção de elipses baseado na transformada de Hough [17].

Essa abordagem consegue segmentar corretamente 99% das faces presentes na base FRGC v2 (3.991 de 4.007). As segmentações classificadas como incorretas se referem àquelas que possuem alguma região que não pertence à face, geralmente cabelo, pescoço ou acessórios para cabeça. Nos experimentos essas imagens são utilizadas sem nenhum pós processamento, para garantir um processo automático [41,42].

Depois que a face foi segmentada, seis pontos de características são detectados e usados para extrair as regiões rígidas e aprimorar o processo de alinhamento. Os pontos detectados são (ver Fig. 4.1a): os cantos internos dos olhos direito e esquerdo, ponta e base do nariz, e os cantos direito e esquerdo da base do nariz. Para realizar essa tarefa nós usamos nossa própria abordagem que detecta corretamente os pontos dos olhos e nariz em 99,92% e 99,6% das imagens da base FRGC v2, respectivamente, como apresentado em [41,42]. Os pontos do nariz não foram corretamente detectados em três imagens devido à ausência do nariz (ver Fig. 4.1b). Os cantos dos olhos não puderam ser detectados corretamente em 16 imagens nas quais a cabeça estava rotacionada (ver Fig. 4.1c). Em nossos experimentos os pontos foram utilizados sem nenhum pós-processamento. No entanto, essas imagens tiveram um impacto negativo mínimo nos nossos resultados devido à nossa abordagem hierárquica que considera diferentes regiões da face (ver Seção 6.1). Maiores detalhes sobre a detecção dos pontos são apresentados em [41,42].

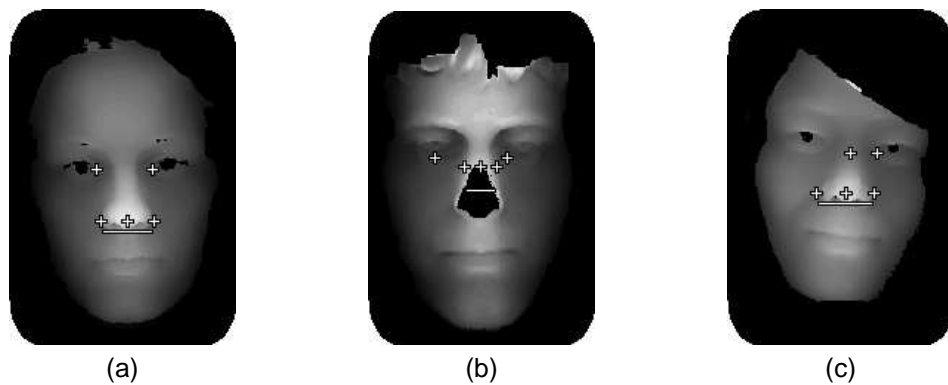


Figura 4.1: Detecção dos pontos característicos: (a) localização dos seis pontos, (b) detecção incorreta dos pontos do nariz, e (c) detecção incorreta dos cantos dos olhos.

Neste trabalho foram consideradas quatro regiões da face (ver Fig. 4.2): (1) região circular e (2) elíptica ao redor do nariz, (3) parte superior da cabeça, incluindo olhos, nariz e testa, e a (4) região da face inteira. As regiões (1) e (2) são utilizadas por sofrerem menor influência das expressões faciais quando comparadas com outras regiões da face [10]. Essas as regiões são as mesmas propostas por Chang *et al.* [10] e podem ser extraídas aplicando funções geométricas nos pontos característicos. As Figs. 4.2a e 4.2b mostram um exemplo dessas regiões.

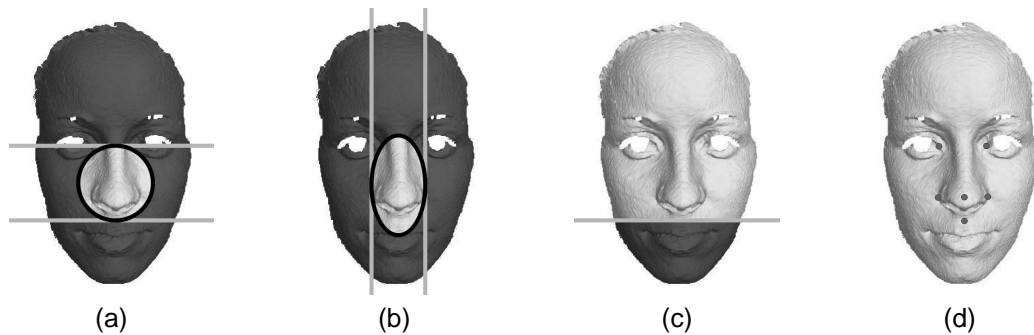


Figura 4.2: Regiões segmentadas da mesma face: (1) região circular e (2) elíptica ao redor do nariz, (3) parte superior da cabeça, e (4) face inteira com os pontos característicos detectados.

Em [5] foi observado que, usando a SIM, a região do nariz não produz resultados discriminatórios o suficiente para o reconhecimento facial. Isso ocorre porque o nariz representa apenas uma pequena fração da face. Por esse motivo, as regiões (3) e (4) também são utilizadas por possuírem mais informações sobre a face inteira. A região (3) é obtida selecionando todos os pontos localizados acima da base do nariz, como apresentado na Fig. 4.2c. Todas as quatro regiões são utilizadas para realizar o reconhecimento facial.

## 5 ALINHAMENTO DE FACES 3D

Nesta seção é apresentado o processo de alinhamento entre duas imagens faciais 3D. O alinhamento das imagens é realizado usando uma abordagem baseada no SA. A SIM é utilizada como medida para avaliar a qualidade e precisão do alinhamento entre as duas superfícies faciais. Maiores detalhes sobre as técnicas utilizadas são apresentadas a seguir.

Em [50] é apresentada uma comparação abrangente entre o ICP, guiado pelo MSE, e uma abordagem estocástica baseada em *Algoritmos Genéticos* (AG), guiada pela SIM, para realizar o alinhamento de faces 3D. Os resultados mostraram que o ICP tende a convergir para soluções de mínimo local devido ao MSE. Por outro lado, ao usar uma abordagem estocástica com a SIM é possível convergir para um mínimo global. Trabalhos anteriores [50, 51] mostraram que a SIM é mais discriminatória que o MSE para avaliar a qualidade do alinhamento entre duas imagens de profundidade. Uma desvantagem do AG é que, embora tenham sido obtidos resultados mais precisos, ele possui um alto tempo de execução.

Em [5] a abordagem baseada em AG foi substituída pelo SA, que é muito mais rápido que o AG, porém ainda é mais lento que o ICP. Esse trabalho também apresenta uma comparação entre o ICP e o SA, e comprova que o SA produz resultados mais precisos. Além disso, os resultados apresentados em [5] mostraram que a SIM funciona bem até mesmo com alinhamentos produzidos pelo ICP.

No presente trabalho, nós usamos o SA para realizar o alinhamento das imagens de profundidade combinado com duas medidas robustas: (1) MSAC [53], e (2) SIM [50, 51]. O SA inicia o processo de alinhamento com o MSAC e um ajuste fino é calculado com a SIM [45, 47]. O valor da SIM resultante é então utilizado para avaliar a qualidade do alinhamento final entre duas faces 3D. As seções seguintes apresentam uma breve explicação sobre a SIM e o SA.

## 5.1 Surface Interpenetration Measure (SIM)

A SIM foi desenvolvida a partir de uma análise detalhada sobre as características das imagens de profundidade. Foi realizada uma avaliação de duas superfícies alinhadas e renderizadas em cores diferentes, cruzando-as repetidamente na área sobreposta. O efeito da interpenetração resulta da natureza das imagens que apresentam imperfeições causadas por limitações do equipamento de aquisição [14, 51].

Devido a esses fatores, até mesmo superfícies lisas (planas) apresentam alguma “rugosidade”, e por isso podemos assumir que independente das formas das superfícies a interpenetração sempre ocorrerá na região de sobreposição de duas vistas diferentes, obtidas pelo mesmo equipamento. Foi observado também que duas vistas adquiridas da mesma superfície do objeto com a mesma posição e parâmetros do equipamento de aquisição geram duas imagens diferentes, ou seja com superfícies apresentando rugosidades diferentes [51].

Quantificando a interpenetração, é possível guiar o processo de registro para a obtenção de alinhamentos mais precisos assim como avaliar a qualidade do alinhamento final obtido. Para isto foi desenvolvida a seguinte medida baseada no vetor normal à superfície, computado pelo ajuste dos quadrados mínimos em cada ponto [50, 51]. Depois do alinhamento das duas imagens,  $A$  e  $B$ , é identificado um conjunto de pontos interpenetrados em  $A$  com seus respectivos correspondentes em  $B$ . Para cada ponto  $p \in A$  é definida uma vizinhança  $N_p$  para ser uma janela pequena  $n \times n$  centrada em  $p$ .

Com  $q$  denotando um ponto na vizinhança  $N_p$ ,  $c$  o ponto correspondente de  $p$  na imagem  $B$  e  $\vec{n}_c$  o vetor normal a superfície em  $c$ , é definido um conjunto de pontos conforme a expressão abaixo:

$$C_{(A,B)} = \{p \in A \mid [(\overrightarrow{q_i - c}) \cdot \vec{n}_c][(\overrightarrow{q_j - c}) \cdot \vec{n}_c] < 0\} \quad (5.1)$$

onde  $q_i, q_j \in N_p$  e  $i \neq j$ . Este conjunto é composto pelos pontos de  $A$  cujos vizinhos incluem ao menos um par de pontos separados pelo plano tangente, computado a partir dos seus correspondentes em  $B$ . A Fig. 5.1 ilustra a situação acima descrita.

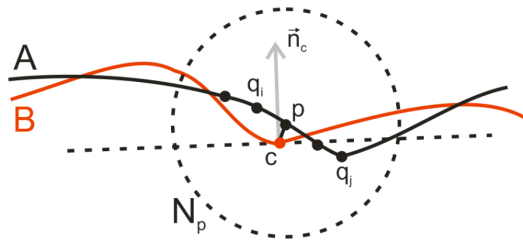


Figura 5.1: Ponto interpenetrado  $p$  de  $A$  em relação a  $B$  [51].

Desta forma, define-se a SIM como a fração de pontos de  $A$  interpenetrados em  $B$ :

$$SIM_{(A,B)} = \frac{|C_{(A,B)}|}{|A|} \quad (5.2)$$

O pseudo-código da SIM é apresentado no Algoritmo 1. Como sugerido em [50], foram aplicadas algumas restrições para o cálculo da SIM com o objetivo de evitar a correspondência imprecisa entre pontos, e obter alinhamentos mais precisos. Foi incluída a restrição  $m = 5^\circ$  como sendo o maior ângulo permitido entre os vetores normais em  $c$  e  $p$ ,  $\vec{n}_c$  e  $\vec{n}_p$ , respectivamente. Definimos que  $p \in C$  se e somente se  $\cos^{-1}(\vec{n}_c \cdot \vec{n}_p) \leq m$ . Além disso, foi utilizada uma restrição para eliminar pontos correspondentes nas regiões de borda da imagem. Neste caso  $p \in C$  se  $c \notin D$ , sendo  $D$  o conjunto de pontos de borda em  $B$ , e a largura da borda definida como  $b = 1$ . Com essas restrições aplicadas à Eq. 5.2, o valor da SIM se mostrou mais eficiente e discriminatório para este problema de reconhecimento de faces 3D.

---

**Algoritmo 1** Surface Interpenetration Measure – SIM.

---

**Entrada:** Duas imagens de profundidade  $A$  e  $B$

- 1:  $C_{(A,B)} \leftarrow \emptyset$ ;
  - 2: **Para**  $p \in A$  **faça**
  - 3:     Define vizinhança  $N_p$ ,  $n \times n$ , com centro em  $p$ ;
  - 4:     Procura o ponto correspondente  $c$  de  $p$  em  $B$ , com  $c \notin D$ ;
  - 5:     Calcula o ângulo  $\theta$  entre os vetores normais  $\vec{n}_p$  e  $\vec{n}_c$ , com relação aos pontos  $p$  e  $c$ ;
  - 6:     **Se**  $\theta < m$  **e**  $\exists \{q_i, q_j\} \in N_p$ , com  $q_i \neq q_j$ , tal que  $q_i, q_j$  estão dispostos em lados opostos do plano tangente em  $c$  **então**
  - 7:          $C_{(A,B)} \leftarrow C_{(A,B)} \cup p$ ;
  - 8:     **Fim Se**
  - 9: **Fim Para**
  - 10: **Retorna**  $\frac{|C_{(A,B)}|}{|A|}$
- 

Observamos em nossos experimentos que o registro entre duas vistas que apresentam uma boa área de interpenetração possuem altos valores da SIM. Os resultados experimentais comparativos mostraram que alinhamentos errados produzem baixos valores da SIM e que pequenas diferenças no MSE podem acarretar grandes diferenças na SIM. Geralmente, um alinhamento preciso produz alto valor para SIM e baixo valor para o MSE.

## 5.2 *Simulated Annealing* (SA)

O SA é um algoritmo de busca local que a partir de uma solução candidata, procura iterativamente uma solução vizinha que melhor resolve o problema [31]. A principal diferença do SA para outros algoritmos de busca local, como por exemplo o *Hill Climbing*, é a possibilidade que o SA tem de aceitar uma solução pior que a atual (solução candidata) durante o processo iterativo. Isso faz com que o SA não fique “preso” em mínimos locais e consiga atingir seu objetivo, que é uma solução próxima ao mínimo global.

*Annealing* é o processo de aquecimento de metais seguido do seu resfriamento lento e gradual com o objetivo de tornar o material mais rígido. Dessa maneira, o SA explora a analogia entre o resfriamento de metais em uma estrutura cristalina de energia mínima com a busca por um mínimo em um sistema qualquer [31]. A Tabela 5.1 apresenta uma relação entre os elementos do processo *Annealing* real e o *Simulated Annealing* [15].

Tabela 5.1: Relação entre o processo *Annealing* real e o *Simulated Annealing* [15].

Annealing	Simulated Annealing
Estados do sistema	Soluções Possíveis
Energia	Custo
Mudança de estado	Soluções vizinhas
Temperatura	Parâmetros de controle
Estado Congelado	Solução Heurística

Para que seja possível resolver um problema utilizando o SA, devem ser definidas duas funções principais:

- a) **Função de custo:** também conhecida como função de objetivo, avalia as soluções do problema indicando o quanto elas o satisfazem.
- b) **Função de incremento:** retorna uma solução vizinha a partir de uma dada solução, permitindo explorar a vizinhança de uma solução candidata.

O algoritmo do SA é apresentado no Algoritmo 2. Inicialmente, o SA recebe como parâmetro uma solução inicial  $x$  e uma variável é definida como sendo a “temperatura” do sistema. Essa temperatura é que vai determinar se uma solução pior deverá ou não ser aceita. Para cada nível da temperatura, uma solução vizinha à solução candidata é criada e avaliada por uma função de custo. Se o custo da solução vizinha for menor, esta passa a ser a solução

candidata. Caso contrário, solução vizinha será aceita com uma probabilidade que segue a distribuição de Boltzman [15,28]. A distribuição de Boltzman utiliza uma relação entre a função de custo da solução atual e a temperatura. É essa probabilidade de aceitar uma solução pior que a atual que permite que o SA não fique “preso” em mínimos locais. A temperatura é decrementada e o processo é executado até que temperatura final  $t_f$  do sistema seja atingida.

---

**Algoritmo 2** Simulated Annealing – SA.

---

**Entrada:** Solução candidata  $x$

```

1:  $x_{melhor} = x$ 
2:  $t = t_0$ 
3: Enquanto  $t > t_f$  faça
4:    $x_v = f_{incremento}(x)$ 
5:    $\Delta E = f_{custo}(x_v) - f_{custo}(x)$ 
6:   Se  $\Delta E \leq 0$  então
7:      $x = x_v$ 
8:     Se  $f_{custo}(x_v) < f_{custo}(x_{melhor})$  então
9:        $x_{melhor} = x_v$ 
10:    Fim Se
11:  Senão
12:     $x = x_v$  com probabilidade  $e^{\frac{-\Delta E}{t}}$ 
13:  Fim Se
14:  Decrementa a temperatura  $t$ 
15: Fim Enquanto
16: Retorna  $x_{melhor}$ 

```

---

Inicialmente, o SA requer que o problema seja modelado de modo que permita a movimentação de uma solução candidata para qualquer outra solução vizinha. Para o alinhamento de imagens 3D são utilizados seis parâmetros:  $T_x$ ,  $T_y$  e  $T_z$  representando translação nos eixos  $X$ ,  $Y$  e  $Z$ , respectivamente;  $R_x$ ,  $R_y$  e  $R_z$  indicando rotação nos eixos  $X$ ,  $Y$  e  $Z$ , respectivamente. Esses parâmetros são usados para compor um vetor de transformação que, aplicado em uma imagem, alinha esta com a outra [15,45]. Neste trabalho o problema de escala entre as imagens não é tratado. Como as imagens faciais 3D utilizadas foram capturadas em um ambiente controlado, não existe grande variação no tamanho das faces. No entanto, se for necessário, é possível adicionar um parâmetros  $S$  no vetor de transformação para ajustar a escala da imagem durante o alinhamento.

A nossa abordagem para o alinhamento possui três etapas principais (ver Fig. 5.2): (1) solução inicial, (2) alinhamento grosseiro e (3) alinhamento fino. Uma solução inicial é obtida alinhando duas imagens pelo centro de massa. Este procedimento transforma as duas imagens para o mesmo centro de coordenadas, otimizando o processo de convergência. Para obter um alinhamento grosseiro, a função de custo do SA minimiza a medida robusta MSAC [53],

que é combinada com o MSE dos pontos correspondentes entre as duas imagens. O MSAC define um limiar que classifica os pontos como *inliers* ou *outliers*. *Inliers* são os pontos que possuem correspondência considerada válida, seguindo algum critério; e *outliers* são os pontos que possuem correspondência inválida [15].

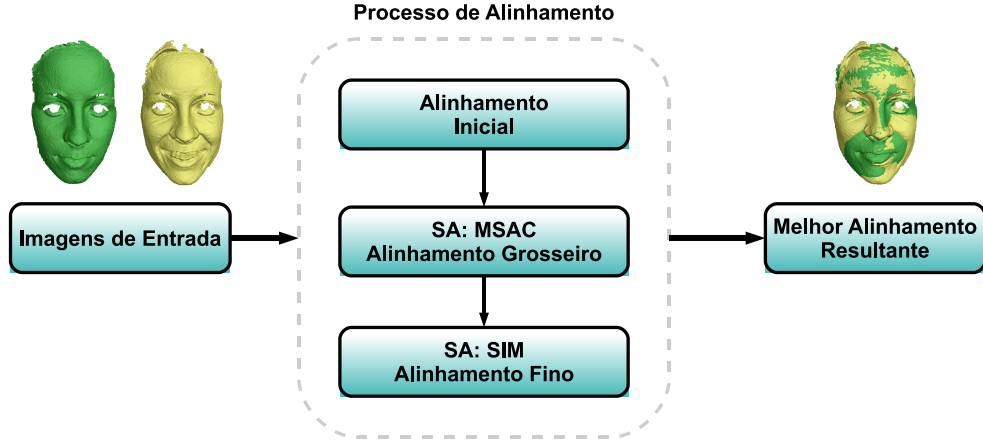


Figura 5.2: Etapas principais do SA: (1) solução inicial, (2) alinhamento grosseiro usando MSAC, e (3) alinhamento fino usando a SIM.

No MSAC, os pontos definidos como *outliers* recebem uma penalidade fixa, enquanto os pontos classificados como *inliers* utilizam a própria medida de erro. Dessa maneira, com a redução do erro associado aos *outliers* é possível diminuir a sua influência durante o processo de alinhamento, e por consequência, obter um alinhamento mais preciso. A Equação 5.3 apresenta a configuração do MSAC, onde,  $\rho$  é o custo de um par de pontos correspondentes,  $e$  é o erro associado ao par de pontos e  $T$  é um limiar escolhido [15].

$$\rho(e^2) = \begin{cases} e^2 & \text{se } e^2 < T^2 \\ T^2 & \text{se } e^2 \geq T^2 \end{cases} \quad (5.3)$$

Na última etapa, o alinhamento fino é obtido usando a SIM como medida de avaliação na função de custo do SA, que tem como objetivo maximizar o número de pontos interpenetrados entre duas superfícies. Detalhes sobre a SIM são apresentados na Seção 5.1.

A temperatura inicial  $t_0$  deve ser “quente” o suficiente para aceitar quase todas as soluções vizinhas da solução candidata. Se esse valor for muito baixo, a solução final será bem próxima da solução inicial. Nesse caso a implementação seria praticamente idêntica a do *Hill Climbing*. Entretanto, se esse valor for muito elevado, o procedimento de busca será praticamente aleatório nas primeiras iterações do SA. A temperatura inicial foi definida empiricamente como  $t_0 = 0,002$  e  $t_0 = 0,07$  para os estágios (2) e (3), respectivamente. Usando



esses valores para  $t_0$ , aproximadamente 60% das soluções ruins são aceitas. De acordo com Rayward-Smith *et al.* [48] este critério deve ser satisfeito para a temperatura inicial.

Na função de incremento do SA, são adicionados a cada elemento do vetor de transformação valores no intervalo  $[-1, 1]$ , para que a melhor solução vizinha seja alcançada. O critério de parada do algoritmo é definido por duas situações: (1) a temperatura final  $t_f$  é atingida, ou (2) a melhor solução não se altera em  $k$  iterações, *i.e.*, o sistema está congelado [36]. O limiar do MSAC foi definido empiricamente como  $T = 3,0$ , utilizando um pequeno conjunto de imagens da base FRGC v1, e depois validados em um conjunto maior de imagens presentes na base FRGC v2.

Foi observado que para produzir um bom alinhamento não é necessário que todos os pontos válidos da imagem sejam utilizados. Uma amostra  $s$  de pontos igualmente espaçados é utilizada. Em nossos experimentos definimos  $s = 20$  e  $s = 30$ , para os estágios (2) e (3), respectivamente, e definimos  $k = 200$ . Embora seja usada uma amostra dos pontos, a SIM que avalia a qualidade final do alinhamento ainda é calculada considerando todos os pontos de ambas as superfícies.

### 5.3 *Simulated Annealing* Modificado

Observamos que alinhamentos entre uma face neutra e outras com expressão podem produzir alinhamentos incorretos. Esses alinhamentos produzem um valor baixo de SIM que, conseqüentemente, prejudica o processo de autenticação. Por esse motivo, e porque a expressão alegre é muito comum quando as pessoas tiram fotos, apresentamos uma abordagem modificada do SA para melhorar os resultados nessas situações. Foi observado que, nas faces com expressões alegres, as regiões mais deformadas são as áreas da boca e das bochechas.

A idéia dessa abordagem é guiar o alinhamento entre uma face neutra e outra com expressão para as áreas mais invariantes às expressões faciais (*e.g.* região no nariz e da testa). Outros autores empregam abordagens semelhantes. Martinez [37] propõe um método para reconhecimento facial 2D que lida com o problema de oclusão, segmentando a face em várias regiões e analisando cada uma delas individualmente. Faltemier *et al.* [21] segmentam a face em 28 regiões, as quais são alinhadas com as suas correspondentes. Então, as regiões que apresentaram menos alterações são combinadas para produzir a medida de autenticação.

Durante o processo de registro do SA, o alinhamento é guiado para essas áreas invariantes da face. Inicialmente, a face é dividida em nove setores, usando os pontos característicos detectados no final da etapa de pré-processamento (ver Seção 4). A Fig. 5.3 ilustra os setores

utilizados e as áreas consideradas invariantes para o processo de alinhamento. No processo original de registro, a SIM é computada com base no número de pontos interpenetrados da superfície  $A$  com relação à  $B$ . Nessa versão modificada, cada ponto interpenetrado  $q$ , que pertence a uma das regiões invariantes, recebe um peso alto (*i.e.*  $w = 500$ ), caso contrário, recebe uma penalidade  $w = 0$ . Em nossos experimentos, observamos que o algoritmo possui bons resultados para outros valores de  $w$ . Ao final do processo de alinhamento, o valor final da SIM é calculado atribuindo peso  $w = 1$  para todos os pontos interpenetrados.

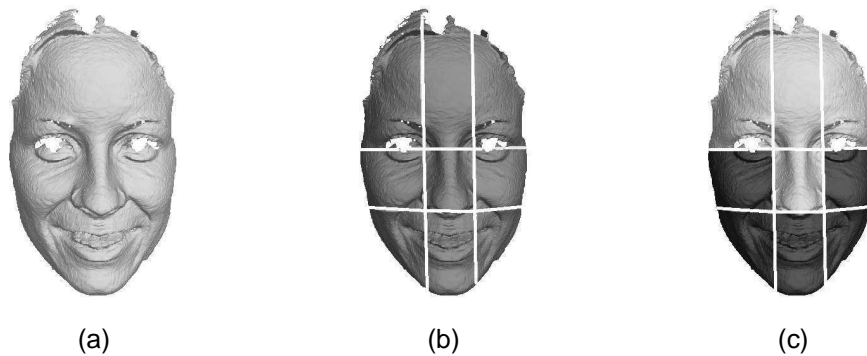


Figura 5.3: Setores utilizados pelo SA modificado: (a) imagem original, (b) nove setores utilizados, e (c) regiões consideradas invariantes em destaque.

A abordagem baseada no SA modificado é indicada para a comparação de uma face neutra com outra que possui expressão facial. Essa técnica pode ser aplicada depois que alinhamento entre duas faces falhar durante o processo de reconhecimento. Isso pode ser realizado usando um modelo de avaliação hierárquico [16]. Os resultados usando esta abordagem podem ser verificados na Seção 6.

## 5.4 Lista de parâmetros

Os parâmetros discutidos nesta seção foram obtidos empiricamente, usando a base FRGC v1 ou subconjuntos da base FRGC v2. Isto implica que o desempenho do sistema é dependente das características da imagem de entrada. Se forem utilizadas imagens adquiridas com outras resoluções, ou outro tipo de sensor ou dispositivos, pode ser que estes parâmetros não sejam aplicáveis. A Tabela 5.2 lista os parâmetros que são utilizados e os valores definidos. Cabe ressaltar que todos os experimentos foram executados utilizando este conjunto de parâmetros, sem qualquer ajuste posterior.

Tabela 5.2: Lista dos parâmetros utilizados.

	Parâmetros	Valor	Descrição
SIM	$m$	5	Maior ângulo permitido entre os vetores normais $\vec{n}_c$ e $\vec{n}_p$ .
MSAC	$T$	3	Penalidade dos <i>outliers</i> .
SA	$t_0$ (2)	0,002	Temperatura inicial para o estágio (2).
	$t_0$ (3)	0,07	Temperatura inicial para o estágio (3).
	$s$ (2)	20	Amostragem de pontos para o estágio (2).
	$s$ (3)	30	Amostragem de pontos para o estágio (3).
	$k$	200	Número de iterações congeladas.
SA mod.	$w$	500	Penalidade para os pontos interpenetrados fora das regiões rígidas.

## 6 AUTENTICAÇÃO DE IMAGENS FACIAIS 3D

Nesta seção são discutidas as abordagens estudadas para o processo de autenticação de faces. Aqui será apresentado como os valores da SIM obtidos durante o alinhamento podem ser utilizados para medir a similaridade entre duas imagens faciais.

Na etapa de autenticação é utilizado o valor da SIM, obtido do alinhamento de duas regiões faciais, como medida de similaridade. Se duas imagens pertencem ao mesmo indivíduo, o alinhamento obtém um valor de SIM alto; caso contrário, o valor obtido será baixo. A faixa de valores da SIM varia de 0% a 100%, no entanto, devido às restrições impostas na SIM, descritas na Seção 5.1, o valor máximo dificilmente ultrapassa 50%. Um experimento inicial foi realizado, na base FRGC v2, para verificar qual região da face é melhor para o reconhecimento facial. Neste experimento foram utilizados os parâmetros descritos na Seção 5.

Para isso, foi utilizado o conjunto de imagens neutras e sem artefatos, e cada imagem foi comparada com todas as restantes, com uma FAR de 0%. A Tabela 6.1 apresenta os resultados de verificação para cada região. A primeira coluna descreve a região da face, seguida da taxa de verificação. A última coluna apresenta o número de falsas rejeições, *i.e.*, número de combinações de um mesmo indivíduo que foram rejeitadas.

Tabela 6.1: Taxa de verificação para cada região da face, com FAR de 0%.

<b>Regiões</b>	<b>Taxa de Verificação</b>	<b>Falsa Rejeição</b>
Nariz círculo (C)	87,4%	627
Nariz elipse (E)	89,6%	519
Região superior da face (U)	85,0%	749
Região da face (F)	89,6%	516
Região da face + SA modificado (M)	87,3%	631

Nesses resultados é possível observar que as regiões elíptica do nariz e face intira possuem a maior taxa de verificação. Os resultados demonstram que a abordagem usando o SA modificado obteve uma taxa de reconhecimento inferior à abordagem original do SA. Entre-

tanto, em [45] foi observado que o SA modificado é mais indicado para realizar a comparação entre uma face neutra e outra com expressão facial. Nessas situações foi reportada uma melhoria de 15% com relação à abordagem original. Foi realizado um segundo experimento, no qual, os valores obtidos para cada região da face foram combinados usando a regra da soma [32]. O resultado da soma é então usado como medida de similaridade. Os resultados são apresentados na Tabela 6.2.

Tabela 6.2: Taxa de verificação usando a regra da soma, com FAR de 0%.

<b>Regras de Combinação</b>	<b>Taxa de Verificação</b>	<b>Falsa Rejeição</b>
C+E	92,6%	369
C+U	95,6%	220
C+F	95,5%	225
C+M	96,4%	178
E+U	97,7%	113
E+F	97,7%	112
E+M	97,3%	136
U+F	88,9%	553
U+M	86,7%	654
F+M	89,8%	508
C+E+U	97,9%	105
C+E+F	97,1%	144
C+E+M	97,4%	131
C+U+F	98,5%	76
C+U+M	98,2%	88
C+F+M	98,3%	86
E+U+F	99,1%	44
E+U+M	99,0%	48
E+F+M	98,9%	56
U+F+M	89,0%	547
C+E+U+F	98,9%	56
<b>C+E+U+F+M</b>	<b>99,2%</b>	<b>39</b>

A melhor taxa de reconhecimento foi obtida combinando a SIM de todas as regiões. Outros trabalhos também empregam a regra da soma combinada com outras medidas para reali-

zar o reconhecimento facial [10,21,35]. De fato, Kittler *et al.* [32] provou que a regra da soma é mais robusta que outros tipos de classificadores. No entanto, outras abordagens podem ser aplicadas para combinação de classificadores, *e.g.*, regra do produto utilizada em [10]. Incluindo na soma a SIM calculada com o SA modificado, o número de falsas rejeições reduz de 56 para 39. Por isso, a métrica  $C + E + U + F + M$  será utilizada como medida de similaridade do nosso método. Essa medida pode ser aplicada em sistemas de verificação e identificação.

## 6.1 Modelo de Avaliação Hierárquico

O modelo de avaliação hierárquico foi proposto para sistemas de verificação que utilizam uma FAR de 0% [16]. Nesta abordagem cada região é analisada em uma hierarquia. As regiões seguintes são analisadas somente se o valor da SIM da região anterior não for discriminatório o suficiente para determinar se as duas pertencem ou não ao mesmo indivíduo.

A classificação é baseada em dois limiares: (1) reconhecimento, e (2) rejeição. Se a SIM obtida for superior ao limiar de reconhecimento, assume-se que ambas as imagens pertencem ao mesmo indivíduo. Se esse valor for inferior ao limiar de rejeição, as imagens são classificadas como sendo de indivíduos distintos. Se a medida de avaliação ficar entre os dois limiares, então não é possível ter uma afirmação, sendo usada a etapa seguinte da hierarquia para tentar uma nova classificação.

Neste trabalho foi usada a mesma abordagem descrita por Drovetto *et al.* [16] para definir os limiares de rejeição e aceitação, onde estes valores são definidos automaticamente para evitar casos de falsa aceitação e rejeição, respectivamente. O limiar de aceitação é definido como sendo o valor da SIM que garante uma falsa aceitação de 0%, sendo então um valor superior à maior SIM computada para faces de mesmos indivíduos. Com isso, o modelo hierárquico é mais adequado para ser empregado em sistemas que requerem 0% de FAR. De maneira análoga, o limiar de rejeição é o menor valor de SIM obtido para combinações entre mesmos indivíduos. Os limiares foram definidos automaticamente para cada um dos subconjuntos de imagens analisados. Outros autores empregam uma abordagem semelhante a esta, não sendo então estabelecido um limiar fixo para cada método [10,20,30,35]. A Fig. 6.1 ilustra como o modelo hierárquico está organizado.

Este método visa alcançar altas taxas de verificação e ao mesmo tempo manter um tempo curto de execução. A ordem na qual cada região é processada é fundamental, pois regiões menores (*e.g.* nariz) são calculadas mais rapidamente que as regiões maiores. Como consequência, o tempo médio para autenticar uma face acaba sendo bem próximo do tempo

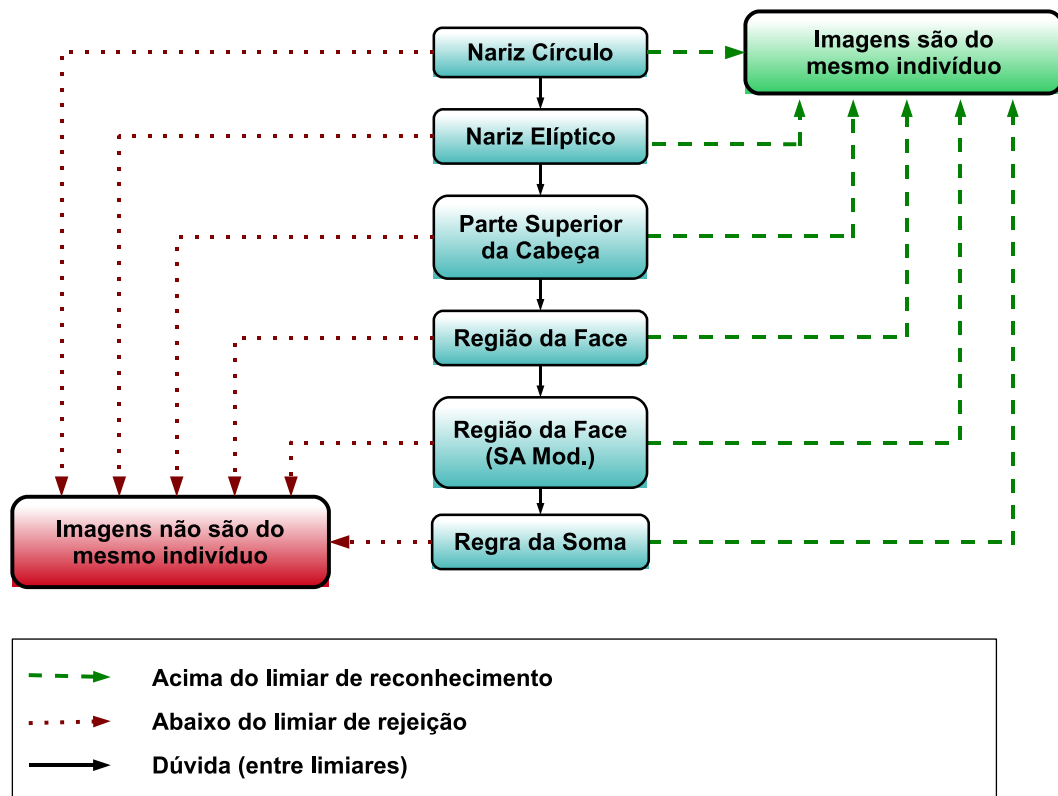


Figura 6.1: Diagrama do Modelo de Avaliação Hierárquico [16].

necessário para alinhar as regiões menores (geralmente cerca de quatro segundos). Na última etapa da hierarquia, é utilizada a regra da soma de todas as regiões. Essa abordagem consegue melhorar a taxa de reconhecimento porque, algumas vezes, uma única imagem consegue alcançar o resultado correto, enquanto todas as outras não. Essa situação pode ser observada em imagens que apresentam oclusão com cabelo, artefatos ou expressão facial. [16].

## 6.2 Modelo de Avaliação Hierárquico Estendido

Neste trabalho também apresentamos uma modificação ao modelo hierárquico. Em cada etapa da hierarquia, ao invés de avaliar apenas a SIM daquele nível, são analisadas também as somas parciais dos níveis já analisados anteriormente. A ordem hierárquica definida é a seguinte: (C) região circular e (E) elíptica do nariz, (U) parte superior da cabeça, (F) região da face inteira e (M) região da face usando o SA modificado, que induz o alinhamento às regiões invariantes da face (*e.g.* testa e nariz) [45,47]. Ao final da hierarquia a soma de todas as regiões é usada para verificar se as duas imagens pertencem ou não à mesma pessoa. Ao todo, 28 medidas são utilizadas e a soma da SIM de todas as regiões é a última medida a ser avaliada. A Tabela 6.3 apresenta as métricas avaliadas em cada nível.

Tabela 6.3: Medidas calculadas em cada nível da hierarquia.

<b>Hierarquia</b>	<b>Medidas de Avaliação</b>
<b>C</b>	$C$
<b>E</b>	$E, C+E$
<b>U</b>	$U, C+U, E+U, C+E+U$
<b>F</b>	$F, C+F, E+F, U+F, C+E+F, C+U+F, E+U+F, C+E+U+F$
<b>M</b>	$M, C+M, E+M, U+M, F+M, C+E+M, C+U+M, C+F+M, E+U+M, E+F+M, U+F+M, C+E+U+M, C+E+U+F+M$



## 7 RESULTADOS EXPERIMENTAIS

Os experimentos foram realizados visando reproduzir um sistema de reconhecimento facial autêntico, no qual é realizado o alinhamento entre uma face adquirida com outra previamente armazenada em um banco de dados. Os experimentos avaliaram dois tipos de sistemas de autenticação: verificação e identificação. Nos sistemas de verificação, o objetivo é responder se a pessoa é quem ela diz ser. O desempenho é avaliado medindo a taxa de verificação obtida para uma determinada Taxa de Falsa Aceitação (FAR). Com uma FAR de 0%, é possível garantir que qualquer pessoa não autorizada não seja aceita pelo sistema de autenticação. Os sistemas de identificação respondem qual é a identidade de uma pessoa, comparando a imagem de entrada com todas do banco de dados. O sistema retorna uma fila ordenada (*rank*) com as faces mais similares, com base em alguma medida de similaridade. O desempenho é medido usando o *rank-one*, que é quando a face com maior similaridade é de fato da mesma pessoa [7].

A FRGC v2 foi segmentada em diversos conjuntos controlados, com um nível crescente de dificuldade, *i.e.*, com a adição de imagens com expressão facial e artefatos. A Tabela 7.1 apresenta a descrição de cada conjunto. Cada imagem da base foi comparada com todas as outras restantes, totalizando mais de 40 milhões de combinações. Para isso, as combinações foram divididas em diversos grupos e executadas em paralelo. Para esses experimentos, o rótulo do indivíduo 04643 foi corrigido para 04783 (ver Seção 3).

O objetivo dos experimentos foi verificar o comportamento do método proposto dentro de cada conjunto controlado. Por exemplo, no experimento de verificação *Level 0 vs. Level 0*, o objetivo é verificar como que o método se comporta em um sistema ideal, o qual é composto apenas por imagens com expressão neutra e sem ruído. Outro exemplo, o experimento *Level 0 vs. All* simula um sistema real de reconhecimento facial, no qual uma imagem qualquer de entrada (conjunto *All*) é comparada com imagens previamente cadastradas e que possuem boa qualidade (conjunto *Level 0*). Para cada subconjunto de experimentos os limiares foram definidos automaticamente. Para cada FAR o limiar é definido de acordo com o maior valor da medida de similaridade obtido com a comparação de indivíduos diferentes.

Tabela 7.1: Classificação dos conjuntos de imagens da base FRGC v2. [47]

Conjunto	Descrição	No. de imagens
<i>First</i>	Inclui a primeira imagem de cada indivíduo armazenada na base.	265
<i>Level 0</i>	Imagens neutras e sem artefatos.	933
<i>Level 1</i>	Todas as faces com expressão neutra.	2.182
<i>Level 2</i>	Inclui faces neutras, pouca expressão e sem artefatos.	2.637
<i>Level 3</i>	Inclui todas as imagens com expressão neutra e pouca expressão com e sem artefatos.	2.884
<i>All</i>	Inclui todas as imagens da base de dados.	4.007

## 7.1 Desempenho do SA para o Alinhamento de Faces 3D

Uma das principais preocupações de usar o SA em sistemas de autenticação se refere ao tempo computacional. Nós realizamos um experimento em um ambiente controlado para verificar o tempo computacional do SA. Um conjunto de 200 imagens aleatórias foi utilizado, sendo cada imagem combinada com todas as demais restantes, totalizando 19.900 combinações. Um computador com as seguintes configurações foi utilizado: Linux O.S., Intel Pentium D 3.4GHz, *cache* de 2MB e 1GB de memória RAM. A Tabela 7.2 apresenta o tempo médio de execução do SA com relação ao registro de duas regiões faciais.

Tabela 7.2: Tempo médio de execução do SA para alinhar duas faces 3D.

Regiões	Tempo médio
C	1,5s
E	1,2s
U	2,0s
F	3,2s
M	3,1s

O tempo médio geral do sistema de autenticação proposto é de aproximadamente 15s. Desse tempo, o processo de aquisição das imagens com um sensor Minolta leva aproximadamente 2s. A etapa de pré-processamento é executada na média em 2s, e o tempo médio necessário para toda etapa de alinhamento é de 11s. No entanto, esse tempo pode ser otimizado com a utilização do modelo de avaliação hierárquico como apresentado em [15, 16].

## 7.2 Experimento I: Verificação

Nesta seção são apresentados os resultados obtidos quando simulado um sistema de verificação real. A idéia é que em tal sistema existe uma base com imagens previamente cadastradas, que possuem características controladas. Esse conjunto de imagens é conhecido como *gallery*. Então, cada imagem que seria capturada durante a utilização do sistema, que podem ter sido capturadas em condições não controladas, fazem parte do conjunto *probe*.

Dessa maneira, nesse experimento cada imagem do conjunto *probe* foi comparada com todas as outras imagens do conjunto *gallery*. A Tabela 7.3 apresenta os resultados para a base FRGC v2. As duas primeiras colunas são os conjuntos *gallery* e *probe* usados nos experimentos, respectivamente. A terceira coluna é a taxa de verificação com 0% de FAR, seguido do número de falsas rejeições *FR*. As duas últimas colunas mostram a taxa de verificação com 0,1% de FAR e o número de falsas rejeições *FR*, respectivamente.

Tabela 7.3: Experimento I: Taxa de verificação usando FAR de 0% e FAR de 0,1%.

<b>Gallery</b>	<b>Probe</b>	<b>FAR 0%</b>	<b>FR</b>	<b>FAR 0,1%</b>	<b>FR</b>
<i>Level 0</i>	<i>Level 0</i>	99,2%	39	100,0%	0
<i>Level 0</i>	<i>Level 1</i>	98,9%	71	99,9%	3
<i>Level 0</i>	<i>Level 2</i>	97,2%	216	99,8%	18
<i>Level 0</i>	<i>Level 3</i>	96,5%	294	99,7%	29
<i>Level 0</i>	<i>All</i>	90,7%	1.067	98,5%	175
<i>Level 1</i>	<i>All</i>	83,5%	4.312	98,2%	475
<i>All</i>	<i>All</i>	70,7%	13.736	96,5%	1.648

Com base nesses resultados, pode ser observado que a adição de imagens com expressão facial e artefatos no conjunto *probe* afeta consideravelmente a taxa de verificação para FAR de 0% e 0,1%. Esse comportamento é esperado, pois o processo de alinhamento considera que a face é um objeto rígido, e por isso não é possível lidar com as deformações causadas pelas expressões faciais. Também foi realizado o experimento “*All vs. All*”, sugerido para a base FRGC v2 [44], no qual cada imagem da base é comparada com todas as restantes. Isso resultou em 16.052.042 combinações, das quais 46.920 são comparações entre mesmos indivíduos. Foi obtida uma taxa de verificação de 96,5% com uma FAR de 0,1%. A Fig. 7.1 apresenta a curva FAR correspondente a este experimento.

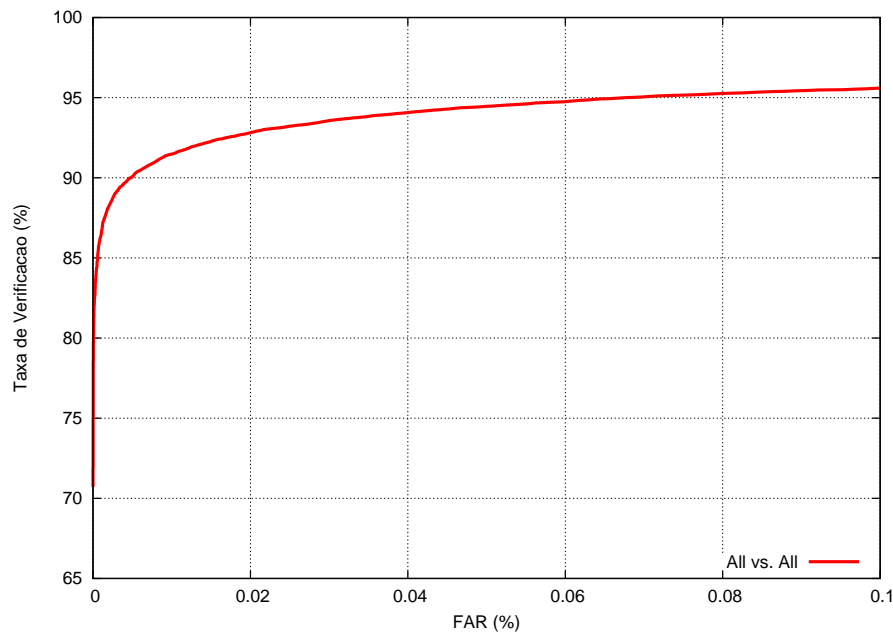


Figura 7.1: Desempenho para o experimento “All vs. All”.

### 7.2.1 Análise dos Buracos na Região dos Olhos

Foi realizado um experimento para verificar se a presença de buracos na região dos olhos afeta o processo de alinhamento. Se os olhos estiverem abertos durante a aquisição das imagens, os buracos são ocasionados pela alta especularidade dos olhos. Se os olhos estiverem fechados, os buracos ocorrem devido à presença dos cílios. Nesse experimento foi usado o conjunto *Level 0*, que foi classificado como: (1) sem e (2) com buracos. A comparação das imagens sem buraco com todas aquelas com buraco totalizou 7.400 combinações.

Foi obtida uma taxa de verificação de 97,4% e 94,4% para a parte superior da cabeça e a face inteira, respectivamente. Para ambas as regiões do nariz a taxa de verificação obtida foi de 100%, usando uma FAR de 0%. Ao utilizar a regra da soma a taxa de verificação se mantém em 100%. De fato, pode ser observado que os buracos na região dos olhos interferem no alinhamento das regiões da face. Essa influência pode ser minimizada com a avaliação individual do nariz, por exemplo, com a utilização do modelo hierárquico (ver Seção 6.1).

### 7.2.2 Análise das Expressões Faciais

Foi realizado um experimento para verificar a influência de cada tipo de expressão facial no processo de reconhecimento. Nesse experimento todas as faces do conjunto *Level 0* foram comparadas com as faces dos conjuntos com expressões, descritos na Tabela 3.2. Os resultados obtidos são apresentados na Tabela 7.4 e Fig. 7.2. A primeira coluna da tabela apre-

sentado o conjunto analisado, seguida pelo número de combinações  $S$  entre mesmos indivíduos e o total de combinações  $TC$ . As duas últimas colunas apresentam a taxa de verificação, com 0% de FAR, e o número de combinações não reconhecidas  $E$ , respectivamente.

Como era esperado, as expressões faciais mais intensas mostraram ser mais difíceis de serem reconhecidas, se comparadas com as de baixa intensidade. Entretanto, pode ser observado que o nosso método é capaz de reconhecer muitas das expressões com uma alta taxa de verificação. As faces com expressão muito nojo e muito alegre foram as mais difíceis de serem reconhecidas, pois elas possuem uma deformação maior na região do nariz. O método alcançou um desempenho baixo porque as regiões segmentadas que são utilizadas incluem o nariz. Faltemier *et al.* [21] também observou em seus experimentos que essas expressões foram as que apresentaram as menores taxas de verificação. As Figs. 3.1g e 3.1h mostram um exemplo das expressões muito alegre e muito nojo, respectivamente.

Tabela 7.4: Resultados para base FRGC usando faces com expressões faciais.

Conjuntos	$S$	$TC$	Verificação (%)	$E$
Pouco alegre	682	194.064	89,4	72
Muito alegre	610	232.317	70,3	181
Pouco surpresa	202	52.248	97,5	5
Muito surpresa	569	160.476	91,9	46
Pouco triste	227	52.248	95,5	10
Muito triste	132	160.476	94,6	7
Pouco boca aberta	135	59.712	96,2	5
Muito boca aberta	588	197.796	95,9	24
Pouco franzir	64	29.856	95,3	3
Muito franzir	188	48.516	86,1	26
Pouco nojo	43	12.129	100,0	0
Muito nojo	82	55.047	54,8	37
“Bochechudo”	442	139.950	84,8	67
Outros	158	61.578	93,6	10

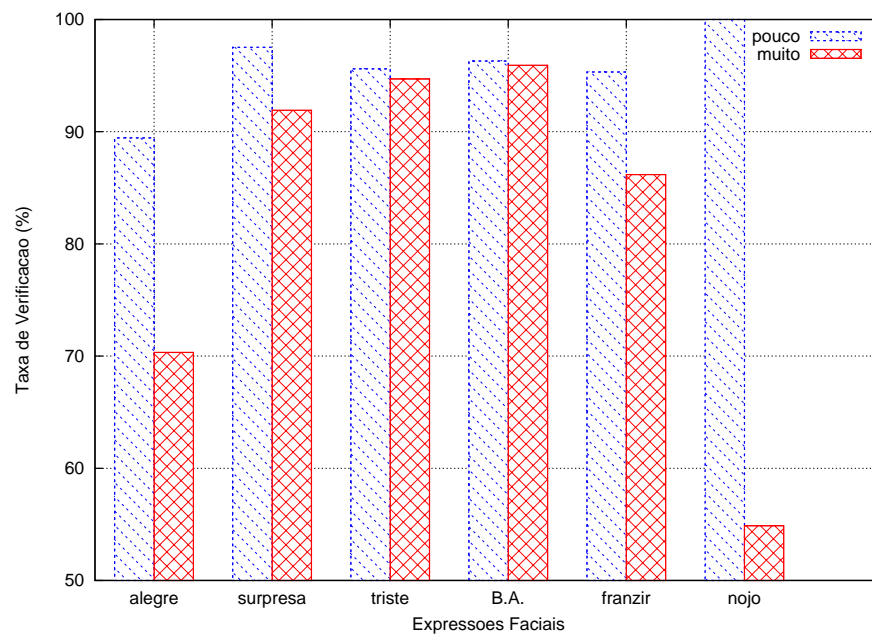


Figura 7.2: Resultados para base FRGC usando faces com expressões faciais.

### 7.2.3 Análise do Modelo Hierárquico

Foi realizado um outro experimento para verificar o comportamento do modelo hierárquico (ver Seção 6.1) e o modelo hierárquico estendido (ver Seção 6.2). Esse experimento foi realizado no modo de verificação considerando uma FAR de 0%. Os resultados obtidos para cada nível do modelo hierárquico e modelo hierárquico estendido são apresentados nas Tabela 7.5 e 7.6, respectivamente. A primeira coluna indica o experimento realizado, e as demais colunas apresentam a taxa de verificação obtida em cada nível da hierarquia.

Tabela 7.5: Taxa de verificação em cada nível no Modelo de Avaliação Hierárquico.

Conjuntos	C	E	U	F	M	Soma
<i>Level 0 vs. Level 0</i>	87,4	92,8	96,0	99,0	99,0	99,3
<i>Level 0 vs. Level 1</i>	86,4	92,0	96,0	98,9	98,9	99,1
<i>Level 0 vs. Level 2</i>	81,7	88,8	93,3	97,3	97,4	97,6
<i>Level 0 vs. Level 3</i>	79,9	87,1	91,9	96,5	96,5	96,9
<i>Level 0 vs. All</i>	69,7	77,7	84,0	90,8	90,9	91,7
<i>Level 1 vs. All</i>	60,2	72,9	81,0	85,9	86,1	86,2
<i>Fall2003 vs. Spring2004</i>	53,4	60,7	68,2	77,9	78,1	79,1
<i>All vs. All</i>	45,7	59,5	68,2	74,9	75,1	75,1

Tabela 7.6: Taxa de verificação em cada nível no Modelo de Hierárquico Estendido.

Conjuntos	C	E	U	F	M	Soma
<i>Level 0 vs. Level 0</i>	87,4	93,6	98,5	99,4	99,4	99,4
<i>Level 0 vs. Level 1</i>	86,4	92,7	98,3	99,1	99,2	99,2
<i>Level 0 vs. Level 2</i>	81,7	89,6	96,5	97,8	97,9	98,0
<i>Level 0 vs. Level 3</i>	79,9	87,9	95,5	97,0	97,3	97,3
<i>Level 0 vs. All</i>	69,7	79,0	89,7	91,8	92,5	92,6
<i>Level 1 vs. All</i>	60,2	73,2	85,0	87,9	88,1	88,1
<i>Fall2003 vs. Spring2004</i>	53,4	62,4	78,0	80,9	82,6	82,6
<i>All vs. All</i>	45,7	59,9	74,1	77,2	77,7	77,7

Com base nesses resultados, é possível observar que mais da metade das faces comparadas foram reconhecidas usando apenas as regiões do nariz. Dessa maneira, as regiões maiores da face, que requerem maior tempo de execução, precisam ser alinhadas em um número reduzido de casos. Com isso o tempo médio para alinhar duas faces acaba ficando muito próximo ao tempo necessário para alinhar as regiões menores [15, 16]. Além disso, a utilização das somas parciais no modelo hierárquico estendido produz maiores taxas de verificação a partir do segundo nível da hierarquia, quando comparado com o modelo hierárquico original.

Os resultados também apresentam que o SA modificado aumenta a taxa de verificação do sistema. É possível verificar na Tabela 7.5, que a utilização do SA modificado apresentou um desempenho melhor nos experimentos que foram utilizadas as imagens com expressões faciais. Além disso, o SA modificado apresentou melhor desempenho quando empregado com o modelo hierárquico estendido.

### 7.3 Experimento II: Identificação

Nesta seção são apresentados os resultados obtidos quando simulado um sistema de identificação real. Nesse experimento uma imagem facial qualquer de entrada é comparada com todas as outras presentes em uma base, criada previamente, conhecida como *gallery*. As imagens utilizadas para simular as faces de entradas pertencem ao conjunto *probe*.

Nesse experimento, quatro conjuntos foram definidos como *gallery*: (1) *Level 0*, com 933 imagens, (2) *Level 1*, com 2.182 imagens, (3) *First* com 465 imagens relativas à primeira imagem de cada indivíduo na base, e (4) *All*, com 4.007 imagens. O conjunto *probe* inclui as

imagens restantes da base que não estão presentes no conjunto *gallery*. Além disso, só foram avaliadas as pessoas do conjunto *probe* que também possuíam uma imagem no conjunto *gallery*. Para o experimento usando o conjunto *All*, cada imagem foi comparada com as 4.006 imagens restantes na base de dados. A curva *Cumulative Match Characteristic* (CMC) é apresentada na Fig. 7.3 e os resultados obtidos no *rank-one* são apresentados na Tabela 7.7. O gráfico CMC apresenta a taxa de identificação acumulada para cada nível do *rank*, que seria o número de imagens do conjunto retornado como mais similares à imagem de entrada [7].

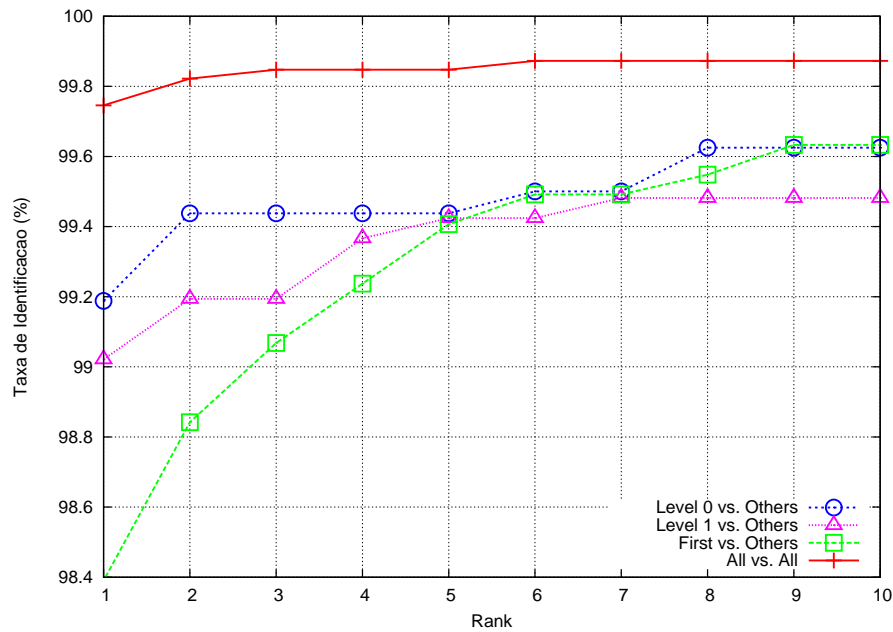


Figura 7.3: Conjunto de curvas CMC: Level 0 vs. Others, Level 1 vs. Others, First vs. Others, e All vs. All.

Tabela 7.7: Experimento II: Taxa de reconhecimento *rank-one*.

Gallery	Probe	Rank-one	Falsa Rej.
<i>Level 0</i>	<i>Others</i>	99,2%	13
<i>Level 1</i>	<i>Others</i>	99,0%	17
<i>First</i>	<i>Others</i>	98,4%	57
<i>All</i>	<i>All</i>	99,7%	10

Devido às suas limitações, o SA pode não ser adequado para sistemas de identificação que utilizem bases de dados muito grandes. Faltemier *et al.* [21] considera aceitável um pequeno atraso na resposta de sistemas de verificação, *e.g.* *check-in* em aeroportos. No entanto, para os sistemas de identificação, que requerem um número muito maior de comparações, existem opções para lidar com o problema de tempo. Em nossos experimentos, por exemplo, incluímos algumas otimizações para melhorar a velocidade do SA, como a adição da etapa de



pré-alinhamento usando o centro de massa e a utilização de uma amostra dos pontos durante a etapa de alinhamento. Mian *et al.* [40] propõem um classificador de rejeição que elimina 97% das imagens da base de dados. Com isso, a comparação de faces 3D seria aplicada apenas em um pequeno conjunto das imagens.

Com base nos resultados experimentais, é possível notar que o método proposto obteve um *rank-one* superior a 98%. O experimento *All vs. All* apresentou a maior taxa de reconhecimento porque um indivíduo pode ter mais de uma imagem armazenada no conjunto *gallery*. De fato, Faltemier *et al.* [21] mostraram que usando uma *gallery* com múltiplas imagens de um mesmo indivíduo é possível obter um melhor desempenho de reconhecimento.

## 7.4 Experimento III: FRGC ROC III

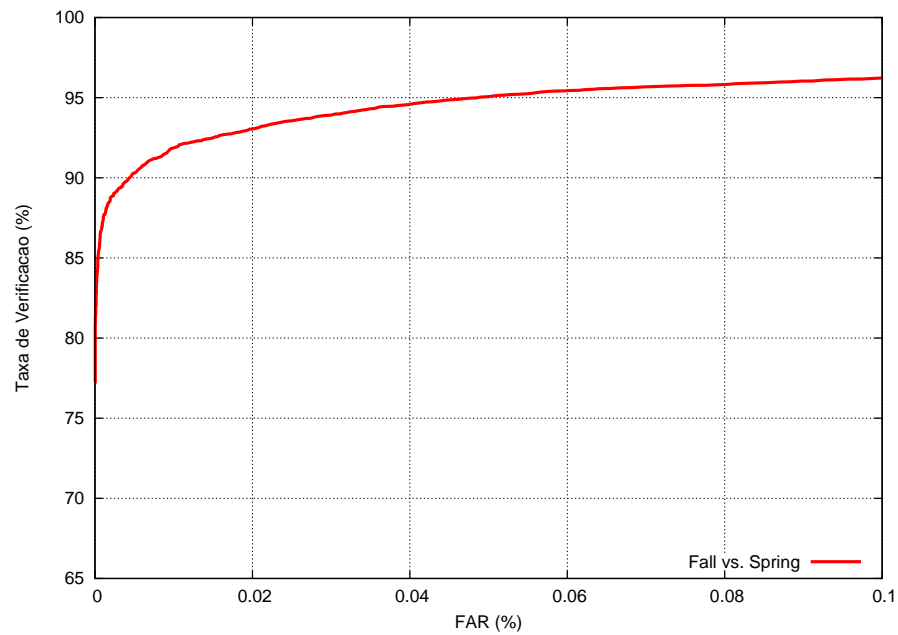
Este experimento é sugerido pelo FRGC como descrito em [44]. O conjunto *gallery* foi criado no outono de 2003 e o conjunto *probe* na primavera de 2004. O intervalo de tempo entre os dois conjuntos aumenta a dificuldade desse experimento. Os resultados para verificação e identificação são apresentados na Tabela 7.8. No experimento de verificação, usando 0% de FAR, foi obtida uma taxa de verificação de 77,2%, enquanto no experimento de identificação foi obtido um *rank-one* de 99,6%. A Fig. 7.4 apresenta as curvas de FAR e CMC.

Tabela 7.8: Experimento III: Resultados para FRGC ROC III.

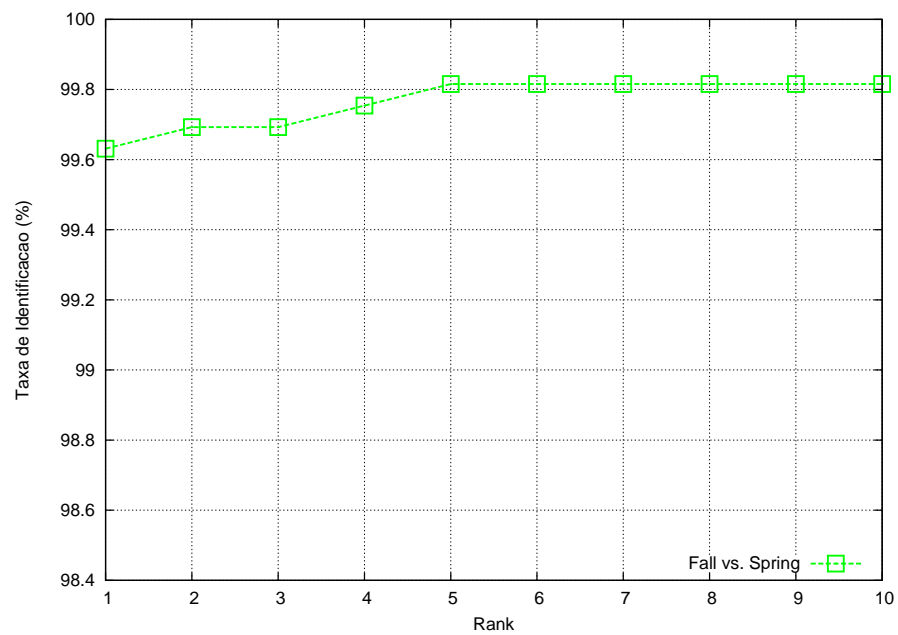
Experimento	Taxa de Reconhecimento
Verificação (FAR 0%)	77,2%
Verificação (FAR 0,1%)	96,6%
Identificação	99,6%

## 7.5 Experimento IV: Modelo de Avaliação Hierárquico

Os resultados dos sistemas de verificação que utilizam uma FAR de 0% podem ser melhorados usando o modelo de avaliação hierárquico para o reconhecimento de faces 3D. A Tabela 7.9 apresenta uma comparação de desempenho entre a regra da soma, o modelo de avaliação hierárquico *H* e o modelo hierárquico estendido *EH*, descrito na Seção 6.1.



(a) FAR Curve



(b) CMC Curve

Figura 7.4: Desempenho no experimento FRGC ROC III: (a) curva FAR, e (b) curva CMC.

Com base nesses resultados é possível observar que com o modelo de avaliação hierárquico é possível melhorar o desempenho do reconhecimento. Além disso, a abordagem baseada no modelo de avaliação hierárquico apresentou melhores resultados quando comparada com as outras duas abordagens avaliadas.

Tabela 7.9: Experimento IV: Taxa de verificação usando FAR 0%.

Conjuntos	Soma	$H$ [16]	$EH$
<i>Level 0 vs. Level 0</i>	99,2%	99,3%	99,4%
<i>Level 0 vs. Level 1</i>	98,9%	99,0%	99,2%
<i>Level 0 vs. Level 2</i>	97,2%	97,6%	98,0%
<i>Level 0 vs. Level 3</i>	96,5%	96,9%	97,3%
<i>Level 0 vs. All</i>	90,7%	91,6%	92,4%
<i>Level 1 vs. All</i>	83,5%	86,1%	88,0%
<i>Fall2003 vs. Spring2004</i>	77,2%	78,9%	82,4%
<i>All vs. All</i>	70,7%	77,0%	78,5%

## 7.6 Comparação com Outros Métodos

Alguns autores também publicaram seus resultados usando a base de dados FRGC v2. Nós reproduzimos os mesmos experimentos usando todas as imagens, incluindo aquelas com artefatos, para possibilitar uma comparação de desempenho com os outros métodos. A Tabela 7.10 mostra os resultados de verificação no experimento *All vs. All*. A Tabela 7.11 apresenta o *rank-one* usando o conjunto *First vs. Others* e os resultados de verificação para o experimento FRGC ROC III estão disponíveis na Tabela 7.12.

Tabela 7.10: Resultados de verificação usando a base *All vs. All*, com 0,1% de FAR.

Método	Taxa de Verificação
Mian <i>et al.</i> [40]	86,6%
Maurer <i>et al.</i> [39]	87,0%
Cook <i>et al.</i> [13]	92,3%
Faltemier <i>et al.</i> [21]	93,2%
Nossa abordagem	<b>96,5%</b>

Tabela 7.11: Resultados de identificação usando conjunto First vs. Others.

<b>Método</b>	<b>Rank-one</b>
Cook <i>et al.</i> [13]	92,0%
Kakadiaris <i>et al.</i> [30]	97,0%
Faltemier <i>et al.</i> [21]	97,2%
Nossa abordagem	<b>98,4%</b>

Tabela 7.12: Resultados de verificação para o experimento FRGC ROC III, com 0,1% de FAR.

<b>Método</b>	<b>Taxa de Verificação</b>
Husken <i>et al.</i> [27]	86,9%
Lin <i>et al.</i> [33]	90,0%
Al-Osaimi <i>et al.</i> [2]	94,1%
Faltemier <i>et al.</i> [21]	94,8%
Kakadiaris <i>et al.</i> [30]	<b>97,0%</b>
Nossa abordagem	96,6%

Além desses experimentos, foi realizado o mesmo experimento descrito em [30], no qual são utilizados dois conjuntos de dados, um apenas com faces neutras e outro usando apenas faces com expressões. Esses conjuntos foram obtidos pela classificação fornecida pela base FRGC v2. Os resultados são apresentados na Tabela 7.13.

Tabela 7.13: Resultados do experimento ROC III em conjuntos diferentes, com 0,1% de FAR.

<b>Conjuntos</b>	<b>Kakadiaris <i>et al.</i> [30]</b>	<b>Nossa abordagem</b>
Expressão neutra	98,5%	<b>99,5%</b>
Expressão não-neutra	<b>95,6%</b>	94,8%

Observando os resultados, o nosso método produziu a melhor taxa de reconhecimento nos dois primeiros experimentos, quando comparado com os outros métodos. No conjunto *Fall vs. Spring* o resultado obtido foi bem próximo ao reportado por [30]. Quando são utilizados os conjuntos classificados por expressão, o nosso método possui o melhor desempenho se são usadas somente faces neutras. Uma das razões é que o nosso método realiza o alinhamento de faces considerando que elas são um objeto rígido. Outros autores que utilizam abordagens similares à nossa reportaram as mesmas limitações. Para reduzir o impacto das expressões faciais eles utilizam regiões pequenas e invariantes da face [10,20,40]. Kakadiaris *et al.* [30] consegue lidar

melhor com as expressões porque a sua abordagem inclui modelos deformáveis da face durante o processo de alinhamento, o que melhora a eficiência do seu método. Cabe ressaltar que um sistema de reconhecimento robusto para faces neutras pode ser usado em diversas aplicações práticas, *e.g.*, *check-in* em aeroportos ou segurança de edifícios.

## 7.7 Experimento V: Melhoria na Taxa de Reconhecimento

Uma das restrições utilizadas no SA, para manter um tempo de execução razoável, foi a limitação do número de iterações  $k$  (ver Seção 5.2). Nós realizamos um experimento para verificar se a taxa de reconhecimento poderia ser melhorada utilizando um número maior de iterações no SA. Nesse experimento a SIM foi recalculada usando  $k = 2000$  apenas nas combinações que teriam maiores chances de atingir um valor maior.

Inicialmente, essa abordagem foi aplicada no conjunto de indivíduos diferentes que teriam sido aceitos usando uma FAR de 0,5% nos experimentos *All vs. All* e *Fall vs. Spring*. Depois disso, o conjunto de falsas rejeições foi calculado novamente e o alinhamento usando mais iterações foi aplicado somente nesses casos. Então, a tabela completa de resultados foi atualizada e todos os experimentos de verificação e identificação foram executados novamente. A Tabela 7.14 apresenta os resultados obtidos.

Como pode ser observado, se for permitido ao SA um número maior de iterações é possível melhorar o desempenho de sistemas de autenticação. Por exemplo, no experimento *Fall vs. Spring*, a taxa de verificação obtida é de 97,4%, superior à todos os resultados apresentados na Tabela 7.12. Pode ser afirmado que a SIM apresenta um grande potencial quando utilizada como medida de similaridade no reconhecimento facial 3D.

Tabela 7.14: Experimento V: Melhoria na taxa de reconhecimento.

	<b>Conjunto</b>	<b>Taxa de reconhecimento</b>
<b>Verificação</b> (FAR de 0,1%)	<i>Level 0 vs. Level 0</i>	100,0%
	<i>Level 0 vs. Level 1</i>	99,9%
	<i>Level 0 vs. Level 2</i>	99,9%
	<i>Level 0 vs. Level 3</i>	99,8%
	<i>Level 0 vs. All</i>	98,8%
	<i>Level 1 vs. All</i>	98,3%
	<i>Fall vs. Spring</i>	97,4%
	<i>All vs. All</i>	97,1%
<b>Identificação</b> (rank-one)	<i>Level 0 vs. Others</i>	99,5%
	<i>Level 1 vs. Others</i>	99,6%
	<i>First vs. Others</i>	99,6%
	<i>Fall vs. Spring</i>	99,8%
	<i>All vs. All</i>	99,9%

## 8 CONSIDERAÇÕES FINAIS

Neste trabalho, foi apresentado um método para reconhecimento de faces 3D. Nosso método usa uma abordagem baseada no *Simulated Annealing* para realizar o alinhamento das imagens de profundidade usando a *Surface Interpenetration Measure* como medida de avaliação. Este trabalho teve como foco a etapa de autenticação das imagens faciais 3D, sendo apresentada uma versão melhorada do modelo de avaliação hierárquico. Experimentos exaustivos foram realizados usando a base FRGC v2 para verificar o desempenho do nosso método, e sua comparação com outros trabalhos estado-da-arte presentes na literatura [13, 21, 27, 30, 33, 39, 40].

Com esta nova abordagem é possível distinguir se duas faces com expressão neutra pertencem ou não ao mesmo indivíduo com uma taxa de verificação de 99%, usando uma FAR de 0%. Nos resultados experimentais, foi observado que comparando uma face neutra com outra que apresenta expressão o nosso método possui um desempenho pouco inferior. Quando usadas todas as imagens da base de dados, no experimento *All vs. All*, as faces podem ser reconhecidas com 99,5% de precisão, usando uma FAR de 0,1%.

Além disso, foi desenvolvida uma melhoria no modelo de avaliação hierárquico para sistemas que requerem FAR de 0% [16]. Incluindo as somas parciais durante a análise hierárquica, a taxa de verificação aumenta de 70,5% para 78,5%, no experimento *All vs. All*. No experimento *Fall vs. Spring* obtivemos uma taxa de verificação de 96,4%, com uma FAR de 0,1% e um *rank-one* de 99,3%. Embora o nosso método sofra quando as imagens contêm expressão facial ou ruído, os resultados obtidos são muito próximos aos melhores resultados já apresentados na literatura [30].

Foi realizado um experimento para verificar se a taxa de reconhecimento pode ser melhorada, ignorando a restrição de tempo do SA. Para isso, o alinhamento com SA foi realizado usando um número maior de iterações, e os resultados melhoraram a taxa de reconhecimento em todos os experimentos. Com isso, nós mostramos o potencial da SIM com medida de similaridade para o reconhecimento facial 3D.

No cenário de identificação, o nosso método atinge 99,5% de reconhecimento quando a primeira imagem de um sujeito é comparada com todas as outras restantes na base. Em todos os experimentos realizados no modo de identificação, nosso método obteve *rank-one* superior a 98%. Pelo nosso conhecimento, estes são os melhores resultados de identificação já apresentados usando a base FRGC v2, quando comparados com outros trabalhos na literatura.



## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ABATE, A. F., NAPPI, M., RICCIO, D., AND SABATINO, G. 2D and 3D face recognition: A survey. *Pattern Recognition Letters* 28, 14 (2007), 1885–1906.
- [2] AL-OSAIMI, F., BENNAMOUN, M., AND MIAN, A. An expression deformation approach to non-rigid 3D face recognition. *Int'l Journal of Computer Vision* 81, 3 (2009), 302–316.
- [3] BELHUMEUR, P. N., HESPANHA, J., AND KRIEGMAN, D. J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* 19, 7 (1997), 711–720.
- [4] BELLON, O. R. P., SILVA, L., AND QUEIROLO, C. 3D face matching using the surface interpenetration measure. In *Lecture Notes in Computer Science*, vol. 3617. Springer-Verlag, 2005, pp. 1051–1058.
- [5] BELLON, O. R. P., SILVA, L., QUEIROLO, C., DROVETTO JR., S., AND PAMPLONA SEGUNDO, M. 3D face image registration for face matching guided by the surface interpenetration measure. In *Proc. IEEE Int'l Conf. Image Processing* (2006), pp. 2661–2664.
- [6] BESL, P. J., AND MCKAY, N. D. A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* 14, 2 (1992), 239–256.
- [7] BOLLE, R., CONNELL, J., PANKANTI, S., RATHA, N., AND SENIOR, A. *Guide to Biometrics*. SpringerVerlag, 2003.
- [8] BOWYER, K. W., CHANG, K., AND FLYNN, P. J. A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. *Computer Vision and Image Understanding* 101 (2006), 1–15.
- [9] BRONSTEIN, M., BRONSTEIN, M. M., AND KIMMEL, R. Expression-invariant 3d face recognition. In *Audio and Video-Based Person Authentication* (2003), Lecture Notes in Comp. Science no. 2688, Springer, pp. 62–69
- [10] CHANG, K. I., BOWYER, K. W., AND FLYNN, P. J. Multiple nose region matching for 3D face recognition under varying facial expression. *IEEE Trans. Pattern Anal. Mach. Intell.* 28, 10 (2006), 1695–1700.
- [11] CHEN, Y. AND MEDIONI, G. Object Modeling by Registration of Multiple Range Images. In *Proc. IEEE Conf. on Robotics and Automation*, 3 (1991), 2724–2729.
- [12] COOK, J., CHANDRAN, V., SRIDHARAN, S., AND FOOKES, C. Face recognition from 3d data using iterative closest point algorithm and gaussian mixture models. In *Proc. Int'l Symposium 3D Data Processing, Visualization and Transmission* (2004), pp. 502–509.

- [13] COOK, J., MCCOOL, C., CHANDRAN, V., AND SRIDHARAN, S. Combined 2D/3D face recognition using log-gabor templates. In *Proc. IEEE Int'l Conf. Video and Signal Based Surveillance* (2006), p. 83.
- [14] DALLEY, G., AND FLYNN, P. J. Range image registration: A software platform and empirical evaluation. In *Proc. Int'l Conf. 3-D Digital Imaging and Modeling* (2001), pp. 246–253.
- [15] DROVETTO JR., S. Reconhecimento facial 3D utilizando o simulated annealing com as medidas surface interpenetration measure e M-estimator sample consensus. Dissertação de mestrado, Universidade Federal do Paraná – (UFPR), 2007.
- [16] DROVETTO JR., S., SILVA, L., AND BELLON, O. R. P. Hierarchical evaluation model for 3D face recognition. In *Proc. Int'l Conf. Computer Vision Theory and Applications* (2008), vol. 2, pp. 67–74.
- [17] DUDA, R. O., AND HART, P. E. Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM* 15, 1 (1972), 11–15.
- [18] EKMAN, P. *Emotion in the Human Face*. Cambridge University Press, 1982.
- [19] ELAD, M., TAL, A., AND AR, M. Content based retrieval of VRML objects - an iterative and interactive approach. In *Proc. Eurographics Workshop on Multimedia* (2002), pp. 107–118.
- [20] FALTEMIER, T., BOWYER, K., AND FLYNN, P. Using a multi-instance enrollment representation to improve 3d face recognition. In *Proc. Int'l Conf. Biometrics: Theory, Applications, and Systems* (2007), vol. 1, pp. 1–6.
- [21] FALTEMIER, T., BOWYER, K. W., AND FLYNN, P. J. A region ensemble for 3d face recognition. *IEEE Trans. Inf. Forensics Security* 3, 1 (2008), 62–73.
- [22] FLYNN, P. J., FALTEMIER, T., AND BOWYER, K. W. 3D face recognition. In *Handbook of Biometrics*, A. K. Jain, A. Ross, and P. J. Flynn, Eds. Springer, 2008, ch. 11, pp. 211–229.
- [23] GELFAND, N., IKEMOTO, L., RUSINKIEWICZ, S., AND LEVOY, M. Geometrically stable sampling for the ICP algorithm. In *Proc. Int'l Conf. 3-D Digital Imaging and Modeling* (2003), pp. 260–267.
- [24] GOKBERK, B., AND AKARUN, L. Comparative analysis of decision-level fusion algorithm for 3d face recognition. In *Proc. Int'l Conf. Pattern Recognition* (2006), pp. 20–24.
- [25] HESHER, C., SRIVASTAVA, A., AND ERLEBACHER, G. A novel technique for face recognition using range imaging. In *Proc. Int'l Symposium Signal Processing and Its Applications* (2003), vol. 2, pp. 201–204.
- [26] HUANG, J., HEISELE, B., AND BLANZ, V. Component-based face recognition with 3d morphable models. In *Proc. Int'l Conf. Audio and Video-based Biometric Person Authentication* (2003), pp. 27–34.

- [27] HUSKEN, M., BRAUCKMANN, M., GEHLEN, S., AND DER MALSBURG, C. V. Strategies and benefits of fusion of 2D and 3D face recognition. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (2005), IEEE Computer Society, pp. 174–174.
- [28] INGBER, L. Simulated annealing: Practice versus theory. *Mathl. Comput. Modelling* 18, 11 (1993), 29–57.
- [29] JAIN, A. K., ROSS, A., AND PRABHAKAR, S. An introduction to biometric recognition. *IEEE Trans. Circuits Syst. Video Technol.* 14, 1 (2004), 4–20.
- [30] KAKADIARIS, I., PASSALIS, G., TODERICI, G., MURTUZA, N., AND THEOHARIS, T. Three-dimensional face recognition in the presence of facial expression: An annotated deformable model approach. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 4 (2007), 640–649.
- [31] KIRKPATRICK, S., GELATT, C. D., AND VECCHI, M. P. Optimization by simulated annealing. *Science* 220, 4598 (1983), 671–680.
- [32] KITTLER, J., HATEF, M., DUIN, R., AND MATAS, J. On combining classifiers. *IEEE Trans. Pattern Anal. Mach. Intell.* 20, 3 (1998), 226–239.
- [33] LIN, W.-Y., WONG, K.-C., BOSTON, N., AND HU, Y. H. 3d face recognition under expression variations using similarity metrics fusion. In *Proc. IEEE Int’l Conf. Multimedia and Expo* (2007), pp. 727–730.
- [34] LU, X., AND JAIN, A. K. Deformation modeling for robust 3D face matching. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (2006), vol. 2, pp. 1377–1383.
- [35] LU, X., JAIN, A. K., AND COLBRY, D. Matching 2.5D face scans to 3D models. *IEEE Trans. Pattern Anal. Mach. Intell.* 28, 1 (2006), 31–43.
- [36] LUNDY, M., AND MEES, A. Convergence of an annealing algorithm. *Mathematical Programming: Series A and B* 34, 1 (1986), 111–124.
- [37] MARTINEZ, A. Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 6 (Jun 2002), 748–763.
- [38] MARTINEZ, A. M., AND KAK, A. C. PCA versus LDA. *IEEE Trans. Pattern Anal. Mach. Intell.* 23, 2 (2001), 228–223.
- [39] MAURER, T., GUIGONIS, D., MASLOV, I., PESENTI, B., TSAREGORODTSEV, A., WEST, D., AND MEDIONI, G. Performance of Geometrix Active-ID 3D face recognition engine on the FRGC data. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (2005), IEEE Computer Society, pp. 154–154.
- [40] MIAN, A., BENNAMOUN, M., AND OWENS, R. An efficient multimodal 2D-3D hybrid approach to automatic face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 11 (2007), 1927–1943.
- [41] PAMPLONA SEGUNDO, M., QUEIROLO, C., BELLON, O. R. P., AND SILVA, L. Automatic 3D facial segmentation and landmark detection. In *Proc. Int’l Conf. Image Analysis and Processing* (2007), pp. 431–436.

- [42] PAMPLONA SEGUNDO, M.; BELLON, O.R.P.; SILVA, L.; QUEIROLO, C.C. AUTOMATIC FACE SEGMENTATION AND FACIAL LANDMARK DETECTION IN RANGE IMAGES, *IEEE Trans. Syst., Man, Cybern.* (SUBMETIDO EM 05/06/2009, ACEITO COM "MINOR REVIEW" EM 22/09/2009)
- [43] PAN, G., HAN, S., WU, Z., AND WANG, Y. 3D FACE RECOGNITION USING MAPPED DEPTH IMAGES. IN *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (2005), VOL. 3, PP. 175–175.
- [44] PHILLIPS, P. J., FLYNN, P. J., SCRUGGS, T., BOWYER, K. W., CHANG, J., HOFFMAN, K., MARQUES, J., MIN, J., AND WOREK, W. OVERVIEW OF THE FACE RECOGNITION GRAND CHALLENGE. IN *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (2005), PP. 947–954.
- [45] QUEIROLO, C., PAMPLONA SEGUNDO, M., BELLON, O. R. P., AND SILVA, L. NOISE VERSUS FACIAL EXPRESSION ON 3D FACE RECOGNITION. IN *Proc. Int'l Conf. Image Analysis and Processing* (2007), PP. 171–176.
- [46] QUEIROLO, C., SILVA, L., BELLON, O. R. P., AND PAMPLONA SEGUNDO, M. 3D FACE RECOGNITION USING THE SURFACE INTERPENETRATION MEASURE: A COMPARATIVE EVALUATION ON THE FRGC DATABASE. IN *Proc. Int'l Conf. Pattern Recognition* (2008), VOL. 1.
- [47] QUEIROLO, C. C., SILVA, L., BELLON, O. R. P., AND PAMPLONA SEGUNDO, M. 3D FACE RECOGNITION USING SIMULATED ANNEALING AND THE SURFACE INTERPENETRATION MEASURE. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (TO APPEAR IN 2009).
- [48] RAYWARD-SMITH, V. J., OSMAN, I. H., REEVES, C. R., AND SMITH, G. D. *Modern Heuristic Search Methods*. JOHN WILEY & SONS LTD., 1996.
- [49] RUSINKIEWICZ, S., AND LEVOY, M. EFFICIENT VARIANTS OF THE ICP ALGORITHM. IN *Proc. Int'l Conf. 3-D Digital Imaging and Modeling* (2001), PP. 145–152.
- [50] SILVA, L., BELLON, O. R. P., AND BOYER, K. *Robust Range Image Registration Using Genetic Algorithms and the Surface Interpenetration Measure*, VOL. 60 OF *Machine Perception and Artificial Intelligence*. WORLD SCIENTIFIC PUBLISHING, 2005.
- [51] SILVA, L., BELLON, O. R. P., AND BOYER, K. ROBUST RANGE IMAGE REGISTRATION USING THE SURFACE INTERPENETRATION MEASURE AND ENHANCED GENETIC ALGORITHMS. *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (2005), 762–776.
- [52] SILVA, L., BELLON, O. R. P., GOTARDO, P. F. U., AND BOYER, K. L. RANGE IMAGE REGISTRATION USING ENHANCED GENETIC ALGORITHMS. IN *Proc. IEEE Int'l Conf. Image Processing* (2003), VOL. 2, PP. 711–714.
- [53] TORR, P., AND ZISSERMAN, A. MLESAC: A NEW ROBUST ESTIMATOR WITH APPLICATION TO ESTIMATING IMAGE GEOMETRY. *Computer Vision and Image Understanding* 78 (2000), 138–156.
- [54] ZHAO, W., CHELLAPPA, R., PHILLIPS, P. J., AND ROSENFELD, A. FACE RECOGNITION: A LITERATURE SURVEY. *ACM Computing Surveys* 35, 4 (2003), 399–458.